

GenCore version 5.1.3
Copyright (c) 1993 - 2002 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: December 9, 2002, 16:36:47 ; Search time 13952 Seconds

(without alignments)
8925.667 Million cell updates/sec

Title: US-09-672-725C-1

Perfect score: 4279
Sequence: 1 ggaagcgcagatcgagatgagtg.....caaaaaaaaaaaaaa 4279

Scoring table: IDENTITY_NUC
Gapop 10.0 / Gapext 1.0

Searched: 2054640 seqs, 14551402878 residues

Total number of hits satisfying chosen parameters: 4109280

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database :

Genbank1:*
1: gb_ba:*
2: gb_hlg:*
3: gb_in:*
4: gb_com:*
5: gb_ov:*
6: gb_pat:*
7: gb_ph:*
8: gb_pl:*
9: gb_pr:*
10: gb_ro:*
11: gb_sts:*
12: gb_sy:*
13: gb_un:*
14: gb_vl:*
15: em_ba:*
16: em_fun:*
17: em_hum:*
18: em_in:*
19: em_mu:*
20: em_com:*
21: em_ov:*
22: em_ph:*
23: em_pat:*
24: em_pl:*
25: em_ro:*
26: em_sts:*
27: em_un:*
28: em_vl:*
29: em_vi:*
30: em_hlg_hum:*
31: em_hlg_inv:*
32: em_hlg_other:*
33: em_hlg_mus:*
34: em_hlg_pln:*
35: em_hlg_rod:*
36: em_hlg_mam:*
37: em_hlg_vtc:*
38: em_sy:*
39: em_hlgo_hum:*
40: em_hlgo_mus:*
41: em_hlgo_other:*

score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	4279	100.0	4279	6 AX105057	AX105057 Sequence
2	4277.4	100.0	4279	6 AX105078	AX105078 Sequence
3	4275.8	99.9	4279	6 AX105080	AX105080 Sequence
4	4272.6	99.9	4279	6 AX105082	AX105082 Sequence
5	4221.8	98.7	4317	4 AF045016	AF045016 Canis fam
6	4221.8	98.7	4317	6 AX105059	AX105059 Sequence
7	3847.4	89.4	4045	12 AF269224	AF269224 Synthetic
8	3824	89.4	3934	4 CEFA19568	AF419568 Canis fam
9	3342.8	78.1	4669	6 I08557	I08557 Sequence 3
10	3341.2	78.1	4378	6 E02326	E02326 Multidrug r
11	3333.2	77.9	4669	6 AR091275	AR091275 Sequence
12	3333.2	77.9	4669	6 AR203322	AR203322 Sequence
13	3327.6	77.8	4646	6 AX391099	AX391099 Sequence
14	3324.4	77.7	4646	6 AX336420	AX336420 Sequence
15	3324.4	77.7	4646	6 AX336708	AX336708 Sequence
16	3324.4	77.7	4646	6 I49610	I49610 Sequence 2
17	3324.4	77.7	4646	6 HUMMDR1	M14758 Homo sapien
18	3315.6	77.5	4669	6 AR055785	AR055785 Sequence
19	3311.8	77.4	4186	6 AX108654	AX108654 Sequence
20	3311.2	77.4	4195	6 AX108656	AX108656 Sequence
21	3270.6	76.4	4264	6 AR051647	AR051647 Sequence
22	3270.6	76.4	4264	6 AR051650	AR051650 Sequence
23	3265.6	75.3	4192	9 AF016535	AF016535 Homo sapl
24	3247.6	75.9	3988	6 AX024454	AX024454 Sequence
25	3239.6	75.7	6505	6 AR028671	AR028671 Sequence
26	3239.6	75.7	9318	6 AR028672	AR028672 Sequence
27	3226.6	75.4	8630	6 AX012321	AX012321 Sequence
28	3223.4	75.3	8630	6 AX012320	AX012320 Sequence
29	3202	74.8	3860	6 AX322787	AX322787 Sequence
30	3200.4	74.8	3860	6 AX322789	AX322789 Sequence
31	3190.2	74.6	3840	6 AX481416	AX481416 Sequence
32	3056.8	71.4	3489	4 AB029153	AB029153 Felis cat
33	3001	70.1	3858	4 OA078609	U78609 Ovis aries
34	2989.6	69.9	4296	10 CRUPGP1	M59253 Chinese ham
35	2988	69.8	4304	10 CRUPGP1165	AF257746 Rattus no
36	2949	68.9	4927	10 AF257746	M30597 Mouse multi
37	2946	68.8	4356	10 MUSMDRAA	M3581 Mouse P-gly
38	2943.4	68.8	4924	10 MUSMDRIA	AF286167 Rattus no
39	2942.2	68.8	4323	10 AF286167	AF286167 Rattus no
40	2936.2	68.6	4788	6 AX322793	AX322793 Sequence
41	2846.8	66.5	3987	10 CRUPGP1185	M59254 Chinese ham
42	2787.4	65.1	4280	10 CRUPGP11	M60041 C.grienseus P
43	2724.6	63.7	4298	10 MUSMDR	M14757 Mouse multi
44	2716.6	63.5	4189	6 AX322791	AX322791 Sequence
45	2705	63.2	4233	6 AR123273	AR123273 Sequence

ALIGNMENTS

RESULT 1
AX105057
LOCUS AX105057 4279 bp
DEFINITION Sequence 1 from Patent WO0123540.
ACCESSION AX105057
VERSION AX105057.1 GI:13921209
KEYWORDS
SOURCE dog
ORGANISM Canis familiaris
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Carnivora; Fissipedia; Canidae; Canis.
REFERENCE
1 (bases 1 to 4279)
Stocker,P.J., Steinel-Crespi,D.T., Crespi,C.L., Relf,T.C. and
Patten,C.J.
P-glycoproteins and uses thereof

Pred. No. is the number of results predicted by chance to have a

|||||
Db 1561 CATGAACACCTAATTAATTTGACCTGTGTGAGAGAGAGGCGCCAGCTAGCTGG 1620
Qy 1621 TGGACGAAGACAGAAATCGCCATTGCTGGGCGCTGTGTCCGAACCCAGATTCTCT 1680
Db 1621 TGGACGAAGACAGAAATCGCCATTGCTGGGCGCTGTGTCCGAACCCAGATTCTCT 1680
Qy 1681 GCTGGATGAGCAACGCTCACTGTGACACTGAAAGTGAAGAGTGTTCAGGTGGCCT 1740
Db 1681 GCTGGATGAGCAACGCTCACTGTGACACTGAAAGTGAAGAGTGTGTCCAGGTGGCCT 1740
Qy 1741 GGATAGGCGCAGAAAAAGCGGACCTACCATTTGTATAGCTCATCTTTGTCTAGACTTCG 1800
Db 1741 GGATAGGCGCAGAAAAAGCGGACCTACCATTTGTATAGCTCATCTTTGTCTAGACTTCG 1800
Qy 1801 TAATGCCGATGTCATTCGCTGGTTTGTATGATGAGTCAATTTGGAGAAAGAAATCATGA 1860
Db 1801 TAATGCCGATGTCATTCGCTGGTTTGTATGATGAGTCAATTTGGAGAAAGAAATCATGA 1860
Qy 1861 TGAACATCAGAAAGAGAGGCGATTACTTCAAACTTGTCACAAATGCAGACAAAGAGAA 1920
Db 1861 TGAACATCAGAAAGAGAGGCGATTACTTCAAACTTGTCACAAATGCAGACAAAGAGAA 1920
Qy 1921 TGAATTTGAGTTAGAAAAAGCCTGCTGATCCAAAGTGAAGTGAATGCTTTGGAAT 1980
Db 1921 TGAATTTGAGTTAGAAAAAGCCTGCTGATCCAAAGTGAAGTGAATGCTTTGGAAT 1980
Qy 1981 GTCTCCAAAGATCAGAGGTGAGTTTAAATAAAGAGATCAGCTCCAGAGAGATACA 2040
Db 1981 GTCTCCAAAGATCAGAGGTGAGTTTAAATAAAGAGATCAGCTCCAGAGAGATACA 2040
Qy 2041 TGCACCAAGAGCCAGACAGAAAGCTTGTACAAAAGAGACTTGAATGGAATGTACC 2100
Db 2041 TGCACCAAGAGCCAGACAGAAAGCTTGTACAAAAGAGACTTGAATGGAATGTACC 2100
Qy 2101 TCCAGTTCTCTGTGAGAGATTTCTGAAGCTGAAGCTGAAGTGAAGTGTGTGTGT 2160
Db 2101 TCCAGTTCTCTGTGAGAGATTTCTGAAGCTGAAGCTGAAGTGAAGTGTGTGTGTGT 2160
Qy 2161 TGGATATTTTGTGCTATTTAAACGGAGGCTGCACACGAGCTTTTCAATATATTTTC 2220
Db 2161 TGGATATTTTGTGCTATTTAAACGGAGGCTGCACACGAGCTTTTCAATATATTTTC 2220
Qy 2221 AAGGATTAAGGATCTTTACCCAGATGAGATCCTGAACAAACAGACAGAAATAGTAA 2280
Db 2221 AAGGATTAAGGATCTTTACCCAGATGAGATCCTGAACAAACAGACAGAAATAGTAA 2280
Qy 2281 CATGTTTCTGTATGTTTCTAGTCTCTGGAATTAATTTCTTTTATTAATTTCTCTCA 2340
Db 2281 CATGTTTCTGTATGTTTCTAGTCTCTGGAATTAATTTCTTTTATTAATTTCTCTCA 2340
Qy 2341 GGGCTTCACATTTGGCAAGGCTGGGAGATCTCAGTAAAGGCGCTGCATACATGTTTT 2400
Db 2341 GGGCTTCACATTTGGCAAGGCTGGGAGATCTCAGTAAAGGCGCTGCATACATGTTTT 2400
Qy 2401 CAGATCCATGCTGAGACAGATGTCAGCTGTTGATGACCTTAAACACCACTGAGAC 2460
Db 2401 CAGATCCATGCTGAGACAGATGTCAGCTGTTGATGACCTTAAACACCACTGAGAC 2460
Qy 2461 ATTGAACACAGGCTTGCATGATGCGGCTCAAGTTAAAGGCGCTATAGGTTCCAGGCT 2520
Db 2461 ATTGAACACAGGCTTGCATGATGCGGCTCAAGTTAAAGGCGCTATAGGTTCCAGGCT 2520
Qy 2521 TGCATGATACCCAGAAATATAGCAAACTTGGAGACAGGCTATTAATCTTAATCTA 2580
Db 2521 TGCATGATACCCAGAAATATAGCAAACTTGGAGACAGGCTATTAATCTTAATCTA 2580
Qy 2581 TGGTGGCAATTAACACTTTACTTACCAATTTGATGCAATTTGCAATTTGCAAGAGT 2640
Db 2581 TGGTGGCAATTAACACTTTACTTACCAATTTGATGCAATTTGCAATTTGCAAGAGT 2640
Qy 2641 TGTGAAATGAATGTTGTCTGACAGCACTGAAGAATGAAGAAGAGCTAGAAGAGC 2700
|||||

Db 2641 TGTGAAATGAATGTTGTCTGACAAAGCACTGAAGAATGAAGAAGCTAGAAGAGC 2700
Qy 2701 TGGAGATTTGCTACAGAACGCTACGAAACCTTCCGAACTGTGTTTCTTGACTCGGGA 2760
Db 2701 TGGAGATTTGCTACAGAACGCTACGAAACCTTCCGAACTGTGTTTCTTGACTCGGGA 2760
Qy 2761 GCGAAGTTTGAATCAATGATAGCAGAGTGTGCAAGTACCATCAAGAACTTTGAG 2820
Db 2761 GCGAAGTTTGAATCAATGATAGCAGAGTGTGCAAGTACCATCAAGAACTTTGAG 2820
Qy 2821 GAAAGCACATCTTGGGCTCTCATTTTCTATCACCAAGCAATGATATTTTCTTA 2880
Db 2821 GAAAGCACATCTTGGGCTCTCATTTTCTATCACCAAGCAATGATATTTTCTTA 2880
Qy 2881 TGTGCTGTCTCCGTTTGGTGGCTACTGTGGGCAATAGTATCATGAATCTTCAAGA 2940
Db 2881 TGTGCTGTCTCCGTTTGGTGGCTACTGTGGGCAATAGTATCATGAATCTTCAAGA 2940
Qy 2941 TGTGCTGTGATCTTCCAGCTATTTGCTTGTGCTGCTAGTGGCAATGATGATCTTCAAGA 3000
Db 2941 TGTGCTGTGATCTTCCAGCTATTTGCTTGTGCTGCTAGTGGCAATGATGATCTTCAAGA 3000
Qy 3001 ATTTGCTCTGATTTGACAAAGCCTGATCAGCCACGCTCAAGCCAAATATGTTGAAG 3060
Db 3001 ATTTGCTCTGATTTGACAAAGCCTGATCAGCCACGCTCAAGCCAAATATGTTGAAG 3060
Qy 3061 AAAAGCCCTCTGATTTGACAGCTCAGCCCTCAGGCTCAAGCCAAATATGTTGAAG 3120
Db 3061 AAAAGCCCTCTGATTTGACAGCTCAGCCCTCAGGCTCAAGCCAAATATGTTGAAG 3120
Qy 3121 AAATGACATTTAATGAGGTGCTGTTCACATCTCCATGCAACCAATCCCGCT 3180
Db 3121 AAATGACATTTAATGAGGTGCTGTTCACATCTCCATGCAACCAATCCCGCT 3180
Qy 3181 CCAGGGCTGAGCCCTGAGAGTGAAGAAAGGCGCAGACGCTGACCTGATGATGACCTGG 3240
Db 3181 CCAGGGCTGAGCCCTGAGAGTGAAGAAAGGCGCAGACGCTGATGATGACCTGG 3240
Qy 3241 CTGTGGAGAGACAGATTGTTGAGCTCTGAGAGGCTCTTGTGACCCCTGGCTGGTTC 3300
Db 3241 CTGTGGAGAGACAGATTGTTGAGCTCTGAGAGGCTCTTGTGACCCCTGGCTGGTTC 3300
Qy 3301 AGTGTATTTGATGAGCAAAAGAGATTAAGCACTGAAATGTCAGTGTCCGAGCAACT 3360
Db 3301 AGTGTATTTGATGAGCAAAAGAGATTAAGCACTGAAATGTCAGTGTCCGAGCAACT 3360
Qy 3361 GGGCATGCTGTCTCAGAGCCCATCTGTTTGAATGACCTGACCATTTGCCAGAACTTGCCTA 3420
Db 3361 GGGCATGCTGTCTCAGAGCCCATCTGTTTGAATGACCTGACCATTTGCCAGAACTTGCCTA 3420
Qy 3421 TGGAGACACAGCGGCTGATGACATGAGAGATTTGAGGCAAGGCAAGGAGGCA 3480
Db 3421 TGGAGACACAGCGGCTGATGACATGAGAGATTTGAGGCAAGGCAAGGAGGCA 3480
Qy 3481 CATACACACTTTCATGAGACATCTCCTGAGAAATCAACACCAAGTATGAGCAAGG 3540
Db 3481 CATACACACTTTCATGAGACATCTCCTGAGAAATCAACACCAAGTATGAGCAAGG 3540
Qy 3541 AACCCAGCTCTCTGAGGCGCAAGAAAGCGCATTTGCAATGCTGCTGTTTGAACA 3600
Db 3541 AACCCAGCTCTCTGAGGCGCAAGAAAGCGCATTTGCAATGCTGCTGTTTGAACA 3600
Qy 3601 GCTCATATTTTGTGTTTGGATGAGAGTATCAGTCTGATGATGAGAAATGTAAGAGT 3660
Db 3601 GCTCATATTTTGTGTTTGGATGAGAGTATCAGTCTGATGATGAGAAATGTAAGAGT 3660
Qy 3661 TGTCCAAAGAGCCCTGAGCAAGGCGCAGAGAGGCGCCTCATTTGTGATGCGCCAGC 3720
Db 3661 TGTCCAAAGAGGCGCTGAGCAAGGCGCAGAGAGGCGCCTCATTTGTGATGCGCCAGC 3720
Qy 3721 CTGTGCAACATCCAGAAATGAGATTTAATAGTGTGTTTCAAGATGCAAGTCAAGGA 3780
Db 3721 CTGTGCAACATCCAGAAATGAGATTTAATAGTGTGTTTCAAGATGCAAGTCAAGGA 3780

QY 3781 GCATGCGACACATCAACAGCTGCGCCGAGAAAGCATCTATTTTTCATGTCAGTGT 3840
|||||
Db 3781 GCATGCGACACATCAACAGCTGCGCCGAGAAAGCATCTATTTTTCATGTCAGTGT 3840
QY 3841 CCAGGCTGAGCAAGCGCTAGTGAAGCTGCGCCATATGACCTGTAAATATTTTAT 3900
|||||
Db 3841 CCAGGCTGAGCAAGCGCTAGTGAAGCTGCGCCATATGACCTGTAAATATTTTAT 3900
QY 3901 ATTGTGTTAAACATGCGCTTATCAAGTTAAAGTGAAGCTTATGSAAGT 3960
|||||
Db 3901 ATTGTGTTAAACATGCGCTTATCAAGTTAAAGTGAAGCTTATGSAAGT 3960
QY 3961 ATGTGAAGCTACCTGTTTAAACATCTTCTGCAAGTGAAGTCAACCAAGTTGAG 4020
|||||
Db 3961 ATGTGAAGCTACCTGTTTAAACATCTTCTGCAAGTGAAGTCAACCAAGTTGAG 4020
QY 4021 AGCTTCAGATTTTATATTAAGCAAGCAAGCAAGCAAGTATGATGATTAAGTATC 4080
|||||
Db 4021 AGCTTCAGATTTTATATTAAGCAAGCAAGCAAGCAAGTATGATGATTAAGTATC 4080
QY 4081 TGGTGTATTCATGCTTATTAAGTATTAAGTATTAAGTATTAAGTATTAAGTAT 4140
|||||
Db 4081 TGGTGTATTCATGCTTATTAAGTATTAAGTATTAAGTATTAAGTATTAAGTAT 4140
QY 4141 TATAATTTTGTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTTAT 4200
|||||
Db 4141 TATAATTTTGTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTTAT 4200
QY 4201 GTAAAGATGCTGAATGTTGAATGAATGCTGAATGAATGAATGAATGAATGAATG 4260
|||||
Db 4201 GTAAAGATGCTGAATGTTGAATGAATGCTGAATGAATGAATGAATGAATGAATG 4260
QY 4261 AAAAAAATGCTGAATGTTGAATGAATGCTGAATGAATGAATGAATGAATGAATG 4279
|||||
Db 4261 AAAAAAATGCTGAATGTTGAATGAATGCTGAATGAATGAATGAATGAATGAATG 4279

RESULT 2
AX105078 4279 bp DNA linear PAT 30-APR-2001
LOCUS
DEFINITION Sequence 22 from Patent WO0123540.
ACCESSION AX105078
VERSION AX105078.1 GI:13921228
KEYWORDS
SOURCE
ORGANISM dog;
Canis familiaris
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Carnivora; Fissipedes; Canidae; Canis.
REFERENCE 1 (bases 1 to 4279)
Stockert, P.J., Steimel-Crespi, D.T., Crespi, C.L., Reif, T.C. and
Patten, C.J.
P-glycoproteins and uses thereof
Patent: WO 0123540-A 22 05-APR-2001;
JOURNAL
TITLE
P-glycoproteins and uses thereof
GENE TEST CORPORATION (US)
FEATURES
source
1. 4279
/organism="Canis familiaris"
/db_xref="taxon:9615"
17. 3862
/note="unnamed protein product"
/codon_start=1
/protein_id="CAC37721.1"
/db_xref="GI:13921228"
/translation="MDPEGRKGSAAEKNEKMGKSKNEKKEKPPYSPFAMFRYSN
WDRICMLVGTMAAIIHGAALPLMLVFGMDSPFANGISRNKTPFPIINESTINNT
OHIINLIEEMTYAAYYSIGAGVLAAYIOYSPMCLAAGRLIKRQEPHAIMO
EIQMDVHDYGEIATRELDVSKINSGIKGTOMFQSLAEPFNGEIVGTPRGKLTIL
VILASPVGSLAAWKIISSTFIDKELAIYAKAGVABEYLAIRVYIARFGKKEEL
ERTNKLDEPAKIGIKAKITANISIGAPLLIYASTALAFWGLVSSISYIGVUL
TYEFSVLGASISGASPSIEAFANAGAYEFLKIDNKSIDSYSKSGKPDNKG
NLEFKNVHSYSPSKREVKILKGLNLKVGOSGYVALVNGSGGCTVYQLQGRYLDPD
GMVCIDGODIRITINVRHLREITGVSOEPLVATIEINIRKRENTMDEIKAVKE

AAATDFIMKLENKFDLTVGERGALQSGGOKRIATRALVNPXILLDEATSDTE
SBAVVOVALDARRGRTTIVIAHRLSTVRNADVINGFDGVIVKRGHDELMKEGTY
FKLVMTQJRGNEIELENAIGESSESDALEMSPKSGSLIKRSTRRSIIHAQGGDR
KLGTEJEDNEVPEVSPWRILKLNSTEMEPVAVGJECALINGLQDPESIIIFSR1IGI
FTRDDEPTKRONSNMFSVLVLGIIISITPEFLOGFPFGNAGEILKRLRYVPSM
LRODVSMPDDEKNTTGALITRILANDAOVKAIGSLAVITQNLANKGTIISLITG
WOUTLLAIVPILAIAGVEMKMLSGOALKKRELBGAKATEALENFRVYSLR
BOKFERMAYOSLOYVRNRSRAKHITGVSESTQAMMFEYFAGCFRGAIVANFRM
FODVILFSAIVPGMAVGOVSSFAADYAKAVSAHYIM1IEKSPILDSYSPHLK
NLEGNVFNENVEVNPREDIPVIGLSLEVKKOYLALVSSGCKSTVOLERF
YPLAGSVLIDGKEIKLKNOMLRAHIGVSPILRPSGSLRBNATAYDSRVSHEP
IMQARKENAHIFETIPBKNTNRVGDKTOLSGGOKRIATRALVROPHILLDEA
75ALDSESEKVOEALDKAREGRICIVIAHRLSTIONDLIVFONKVAERHGHQDL
LAKKIIYFSMVSVOGANKR

BASE COUNT 1295 a 833 c 1008 g 1143 t
ORIGIN

Query Match 100.0%; Score 4277.4; DB 6; Length 4279;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 4278; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 GGAGCGGAGGTGGGATGATCTGGAAGAGCGGCTTAAGGAGTGCAGAGAAGACTT 60
|||||
Db 1 GGAGCGGAGGTGGGATGATCTGGAAGAGCGGCTTAAGGAGTGCAGAGAAGACTT 60
QY 61 CTGGAATAATGGGCAAAAAAATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 120
|||||
Db 61 CTGGAATAATGGGCAAAAAAATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 120
QY 121 CACGTTTGCAATGTTTGGCTATTCATTAATGCTGATAGTGTATATGTTGGTGGGAC 180
|||||
Db 121 CACGTTTGCAATGTTTGGCTATTCATTAATGCTGATAGTGTATATGTTGGTGGGAC 180
QY 181 AATGGCTGCCATATCCATGAGAGCTGCACCTCCCTATGATGATCTGTTTGGAAAT 240
|||||
Db 181 AATGGCTGCCATATCCATGAGAGCTGCACCTCCCTATGATGATCTGTTTGGAAAT 240
QY 241 GACAGATAGCTTTGCAAAATGACAGAAATTTCAAGAAACAAATTTCCAGTTATATTA 300
|||||
Db 241 GACAGATAGCTTTGCAAAATGACAGAAATTTCAAGAAACAAATTTCCAGTTATATTA 300
QY 301 TGAAGATTTACAAATATACCAATATTCATCAACATCTGAGAGAGAGAGAGAGAG 360
|||||
Db 301 TGAAGATTTACAAATATACCAATATTCATCAACATCTGAGAGAGAGAGAGAGAG 360
QY 361 GTATGCTTATTTATTAAGTGGATGAGTGGCTGGCTGGCTGGCTGGCTGGCTGGCT 420
|||||
Db 361 GTATGCTTATTTATTAAGTGGATGAGTGGCTGGCTGGCTGGCTGGCTGGCTGGCT 420
QY 421 TTTCAATCTGCTGCTGGCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 480
|||||
Db 421 TTTCAATCTGCTGCTGGCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 480
QY 481 TGCATATCATGAG 540
|||||
Db 481 TGCATATCATGAG 540
QY 541 CCGGCTCAG 600
|||||
Db 541 CCGGCTCAG 600
QY 601 CTTTCACTCAATAGCAACATTTTTCACCGGTTTATAGTGGGTTTACACGTGTGGAA 660
|||||
Db 601 CTTTCACTCAATAGCAACATTTTTCACCGGTTTATAGTGGGTTTACACGTGTGGAA 660
QY 661 GCTAACCTTGATTTTGGCCATCAGCCCTGTTCTGAGACTTTCAGCCCATCTGGGC 720
|||||
Db 661 GCTAACCTTGATTTTGGCCATCAGCCCTGTTCTGAGACTTTCAGCCCATCTGGGC 720
QY 721 AAGATATCTATCTTCAATTAAGTAAAGAGAGAGAGAGAGAGAGAGAGAGAGAG 780
|||||
Db 721 AAGATATCTATCTTCAATTAAGTAAAGAGAGAGAGAGAGAGAGAGAGAGAGAG 780

OY	781	AGCTGAAGAAGTCTTAACCATTCAGAACTGATGGCTTTGGAGACAATAAAGA	840
Db	781	AGCTGAAGAAGTCTTAACCATTCAGAACTGATGGCTTTGGAGACAATAAAGA	840
OY	841	ACTTGAAAGGTACACAAAATTTAGAAGAGCTAAAGCAATTGGGATTAAGAAAGCAT	900
Db	841	ACTTGAAAGGTACACAAAATTTAGAAGAGCTAAAGCAATTGGGATTAAGAAAGCAT	900
OY	901	CAGGCCAACATTTCTATTTGGTCCGCTTCTTATTTGATCTATGCATATGCTCGCC	960
Db	901	CAGGCCAACATTTCTATTTGGTCCGCTTCTTATTTGATCTATGCATATGCTCGCC	960
OY	961	TTTTGGTATGGAGCCTCTGCTGCTCCAGTGAATATCTATTTGACAAGTACAC	1020
Db	961	TTTTGGTATGGAGCCTCTGCTGCTCCAGTGAATATCTATTTGACAAGTACAC	1020
OY	1021	TGTCCTCTTTCTGTATTAATTGGGCTTTTAATTTGATTTGACAGGCATCCCAGCATTTGA	1080
Db	1021	TGTCCTCTTTCTGTATTAATTGGGCTTTTAATTTGATTTGACAGGCATCCCAGCATTTGA	1080
OY	1081	AGCATTTGCAACGACGAGAGGAGAGCTTATGAATCTTCAGAAATTTGACATTAACC	1140
Db	1081	AGCATTTGCAACGAGAGGAGAGCTTATGAATCTTCAGAAATTTGACATTAACC	1140
OY	1141	AAGCATTTGACAGCTATTTGAGAGTGGACATTAACACAGATTAATTTAAGGAATTTGGA	1200
Db	1141	AAGCATTTGACAGCTATTTGAGAGTGGACATTAACACAGATTAATTTAAGGAATTTGGA	1200
OY	1201	ATTCAAAAATGTTCACTTACCTTACCTTCGAAAAAGTTTAAGTCTTTAAAGGCTCT	1260
Db	1201	ATTCAAAAATGTTCACTTACCTTACCTTCGAAAAAGTTTAAGTCTTTAAAGGCTCT	1260
OY	1261	CANACTGAAAGTTCAAGAGTGGGACAGAGTGGCGCTGGTTGGAAACAGTGGCTGGCGGAA	1320
Db	1261	CANACTGAAAGTTCAAGAGTGGGACAGAGTGGCGCTGGTTGGAAACAGTGGCTGGCGGAA	1320
OY	1321	GAGCAGACACCTGCGAGCTGATGACAGAGGCTCTATGACCCACAGATGGCATGCTGTAT	1380
Db	1321	GAGCAGACACCTGCGAGCTGATGACAGAGGCTCTATGACCCACAGATGGCATGCTGTAT	1380
OY	1381	TCATGGACAGGACATTTAGGACCATTAATGTAAGGCATCTTCGGGAAATTACTGCTGTGT	1440
Db	1381	TCATGGACAGGACATTTAGGACCATTAATGTAAGGCATCTTCGGGAAATTACTGCTGTGT	1440
OY	1441	GAGTGAAGAGCCTGCTGTTGTCACACAGATTAAGGAAATTCGCTATGTCGCGCA	1500
Db	1441	GAGTGAAGAGCCTGCTGTTGTCACACAGATTAAGGAAATTCGCTATGTCGCGCA	1500
OY	1501	AAATGTCAACATGATGATGTAAGAAAGCTTTAAGGAAGCCAATGCCATGATTTTAT	1560
Db	1501	AAATGTCAACATGATGATGTAAGAAAGCTTTAAGGAAGCCAATGCCATGATTTTAT	1560
OY	1561	CATGAAATACCTTAATTAATTTGACACTGCTGTTGGAGAGACAGGGCCACGCTGAGTGG	1620
Db	1561	CATGAAATACCTTAATTAATTTGACACTGCTGTTGGAGAGACAGGGCCACGCTGAGTGG	1620
OY	1621	TGGACAGAAACAGAGAAATCCGCAATTTGCTCGGCCCTGTTGCGAAACCCCAAGTTCCTT	1680
Db	1621	TGGACAGAAACAGAGAAATCCGCAATTTGCTCGGCCCTGTTGCGAAACCCCAAGTTCCTT	1680
OY	1681	GCTGGATGAGGACAGGACGCTGCTGACACATGTAAGATGAAGCAGTGGTCCAGTGGCCCT	1740
Db	1681	GCTGGATGAGGACAGGACGCTGCTGACACATGTAAGATGAAGCAGTGGTCCAGTGGCCCT	1740
OY	1741	GGATTAAGGACAGAAAGCCGCACTACCATTTGATGATGCTACGCTTTCTACAGTTTCG	1800
Db	1741	GGATTAAGGACAGAAAGCCGCACTACCATTTGATGATGCTACGCTTTCTACAGTTTCG	1800
OY	1801	TAATGCGGATGATTTGCTGCTGTTTATGATGAGTGCATTTGGAAGAAAGCAATCATGA	1860
Db	1801	TAATGCGGATGATTTGCTGCTGTTTATGATGAGTGCATTTGGAAGAAAGCAATCATGA	1860
OY	1861	TGAATCTATGAAAGAGAGGCACTTACTTCAAACTGTCAATATGACAGACAGGAGAA	1920
Db	1861	TGAATCTATGAAAGAGAGGCACTTACTTCAAACTGTCAATATGACAGACAGGAGAA	1920
OY	1921	TGAATTTGATTTGAAATGCACTGTTGATTCGAAAAGTGAAGTATGCTTGAAT	1980
Db	1921	TGAATTTGATTTGAAATGCACTGTTGATTCGAAAAGTGAAGTATGCTTGAAT	1980
OY	1981	GTCCTCAAAGATTCAGGGTCCAGTTTATTAATAAAGATCACTCGCAGATATACA	2040
Db	1981	GTCCTCAAAGATTCAGGGTCCAGTTTATTAATAAAGATCACTCGCAGATATACA	2040
OY	2041	TGCACCAAGGCCCCAAGACAGAAAGCTTGTATCAAAAGAGGACTTGAATGATATAC	2100
Db	2041	TGCACCAAGGCCCCAAGACAGAAAGCTTGTATCAAAAGAGGACTTGAATGATATAC	2100
OY	2101	TCCAGTTCCCTTCGAGAGATTTGGAAGCTGAACCTCAACTGATATGCTTATTTGGCT	2160
Db	2101	TCCAGTTCCCTTCGAGAGATTTGGAAGCTGAACCTCAACTGATATGCTTATTTGGCT	2160
OY	2161	TGCTATATTTTGCTATTTATTAACGAGAGGCTTCAACACAGATTTTCAATTAATTTTC	2220
Db	2161	TGCTATATTTTGCTATTTATTAACGAGAGGCTTCAACACAGATTTTCAATTAATTTTC	2220
OY	2221	AAGATTTATAGGATCTTTTACCAGAGATGAGATCTTGAAACAAAAGACAGATATGA	2280
Db	2221	AAGATTTATAGGATCTTTTACCAGAGATGAGATCTTGAAACAAAAGACAGATATGA	2280
OY	2281	CATGTTTCTGTATTTCTAGTCCTTGAATTAATTTCTTTATTAACATTTTCCCTCA	2340
Db	2281	CATGTTTCTGTATTTCTAGTCCTTGAATTAATTTCTTTATTAACATTTTCCCTCA	2340
OY	2341	GGGCTTCACTTTGGCAAGAGTGGGAGATCTCACATAGCGGCTTCGATACATGGTTT	2400
Db	2341	GGGCTTCACTTTGGCAAGAGTGGGAGATCTCACATAGCGGCTTCGATACATGGTTT	2400
OY	2401	CAGATCCATGCTGAGACAGATGACAGCTGTTGATGAAACCCGTAACACACACTGGAGC	2460
Db	2401	CAGATCCATGCTGAGACAGATGACAGCTGTTGATGAAACCCGTAACACACACTGGAGC	2460
OY	2461	ATTGACAAACAGGCTTGCCAAATGATGCGGCTCAAGTTAAAGGGCTATAGTTCACAGCT	2520
Db	2461	ATTGACAAACAGGCTTGCCAAATGATGCGGCTCAAGTTAAAGGGCTATAGTTCACAGCT	2520
OY	2521	TGCTGCTATTAACCCAGATTAACCAATCTTGGAAGAGGACTTATTAATCTTAATCTA	2580
Db	2521	TGCTGCTATTAACCCAGATTAACCAATCTTGGAAGAGGACTTATTAATCTTAATCTA	2580
OY	2581	TGCTTGCAATTTAAGCTTTTACTTTAGCAATTTGACCATCATTTGCAATAGCAGAGT	2640
Db	2581	TGCTTGCAATTTAAGCTTTTACTTTAGCAATTTGACCATCATTTGCAATAGCAGAGT	2640
OY	2641	TGTTGAATGAATAATTTGCTGTCGACACACATGAAAGTATGAAGAGCTAGAGAGC	2700
Db	2641	TGTTGAATGAATAATTTGCTGTCGACACACATGAAAGTATGAAGAGCTAGAGAGC	2700
OY	2701	TGGGAAGATTTGCTACAGAAAGCATGCAAAAGTCCGAAGCTGTTGTTCTTGACTCGGGA	2760</

Db	2941	TGTTCTTTTGGTAATTCGACGTAATGTGCTTTGGTGCACATGCGAGTGGGCGAGTCAAGTTC	3000
Qy	3001	ATTTCGCTCGACTATAGCCAAAGCCAAAGATATCAGACGACCACGTCATCATATCAATTGA	3060
Db	3001	ATTTCGCTCGACTATAGCCAAAGCCAAAGATATCAGACGACCACGTCATCATATCAATTGA	3060
Qy	3061	AAAAAGCCTCTGATTTGACAGTACAGCCCTCAGGGCTCAAGCCAAATACGTTGGAAGG	3120
Db	3061	AAAAAGCCTCTGATTTGACAGTACAGCCCTCAGGGCTCAAGCCAAATACGTTGGAAGG	3120
Qy	3121	AAATGTACATTTAATAGAGGTGCTGTTCACATATCCCACTGACACAGACATCCCGTGT	3180
Db	3121	AAATGTACATTTAATAGAGGTGCTGTTCACATATCCCACTGACACAGACATCCCGTGTCT	3180
Qy	3181	CCAGGGGCTGAGCCTCGAGGTAAAGAGGCGCACAGCCTGGCCCTCGTAAGTAGCAAGTGG	3240
Db	3181	CCAGGGGCTGAGCCTCGAGGTAAAGAGGCGCACAGCCTGGCCCTCGTAAGTAGCAAGTGG	3240
Qy	3241	CTGTGGGAAGAAGACAGTTGTTACAGTCCGTAGAGGGCTTCTATGACCCCTTGGCTGTTC	3300
Db	3241	CTGTGGGAAGAAGACAGTTGTTACAGTCCGTAGAGGGCTTCTATGACCCCTTGGCTGTTC	3300
Qy	3301	AGTCTAATTTGATGGCAAAAGATAAAGCACCTGTANTCTCAGTGGCTCCGAGACACCT	3360
Db	3301	AGTCTAATTTGATGGCAAAAGATAAAGCACCTGTANTCTCAGTGGCTCCGAGACACCT	3360
Qy	3361	GGGCAATGCTGTCTCAGAGAGCCATCCCTGTTTACTGACAGATTTCCGAGAAACATTGCTTA	3420
Db	3361	GGGCAATGCTGTCTCAGAGAGCCATCCCTGTTTACTGACAGATTTCCGAGAAACATTGCTTA	3420
Qy	3421	TGAGAGACAACAGCCGGGTCTATCACATGAGAGATTATGCAAGCAGCCAAAGNAGCCAA	3480
Db	3421	TGAGAGACAACAGCCGGGTCTATCACATGAGAGATTATGCAAGCAGCCAAAGGAGCCAA	3480
Qy	3481	CATACACCACTTATTCGAGACACTCCCTGAGAAATACAAACCCAGAGTAGAGACAAGG	3540
Db	3481	CATACACCACTTATTCGAGACACTCCCTGAGAAATACAAACCCAGAGTAGAGACAAGG	3540
Qy	3541	AAACCAGCTCTCTGTTGGGCCCAAGAAAGCCAGTGGCAATAGCTGCGCTCTTGTTAGACA	3600
Db	3541	AAACCAGCTCTCTGTTGGGCCCAAGAAAGCCAGTGGCAATAGCTGCGCTCTTGTTAGACA	3600
Qy	3601	GGCTCATATTTTGGCTTTTGGATGATGAGTACATCAGTCTCGAATACAGAAATGTGAAAAGT	3660
Db	3601	GGCTCATATTTTGGCTTTTGGATGATGAGTACATCAGTCTCGAATACAGAAATGTGAAAAGT	3660
Qy	3661	TGTCACAAGACCCCTGGACAAGCGCAGAGAGGCGCCACCTCATTTGATCGCCACCG	3720
Db	3661	TGTCACAAGACCCCTGGACAAGCGCAGAGAGGCGCCACCTCATTTGATCGCCACCG	3720
Qy	3721	CTTGTTCACCATTCAGAAATGACATTTAATATGTGTGTTTCAGAATGGCAAAAGTCAAGGA	3780
Db	3721	CTTGTTCACCATTCAGAAATGACATTTAATATGTGTGTTTCAGAATGGCAAAAGTCAAGGA	3780
Qy	3781	GGATGGGCACATCAATCAAGAGCTGCTGGCGCCAGAAAGCATCTATTTTTCATAGTGCAGGT	3840
Db	3781	GGATGGGCACATCAATCAAGAGCTGCTGGCGCCAGAAAGCATCTATTTTTCATAGTGCAGGT	3840
Qy	3841	CCAGGCTGAGCAAAAGCGCTAGTAACTGTGGCCATATGAGCTGTAAATATTTTAT	3900
Db	3841	CCAGGCTGAGCAAAAGCGCTAGTAACTGTGGCCATATGAGCTGTAAATATTTTAT	3900
Qy	3901	ATTGTGTTAAACAATGSCATTTTATCAAAAGTTAAAGGAGAGCACTTACTGGAAGAACT	3960
Db	3901	ATTGTGTTAAACAATGSCATTTTATCAAAAGTTAAAGGAGAGCACTTACTGGAAGAACT	3960
Qy	3961	ATGTAGAACTACCTGTTTAAACATTTCTTGTCGAATGAAGATCACTTCCACCAATTTAG	4020
Db	3961	ATGTAGAACTACCTGTTTAAACATTTCTTGTCGAATGAAGATCACTTCCACCAATTTAG	4020
Qy	4021	AGTCTTCAATTTTATATTAATTAAGGAACCAAAAGAAACATTAATCTGATGAATTAATAAC	4080
Db	4021	AGTCTTCAATTTTATATTAATTAAGGAACCAAAAGAAACATTAATCTGATGAATTAATAAC	4080

	99.9%	Score	4275.8;	DB	6;	length	4279;
Query Match	100.0%	Pred.	No. 0;				
Best Local Similarity	0;	Mismatches	2;	Indels	0;	Gaps	0;
Matches 4277; Conservative							

Qy	1	GGAGCCGAGAGTCCGGATGATCTCGAAGAGAGCCGCTAAGGGAGAGTCCAGAGAAACTT	60
Db	1	GGAGCCGAGAGTCCGGATGATCTCGAAGAGAGCCGCTAAGGGAGAGTCCAGAGAAACTT	60
Qy	61	CTGAAATGSGCCAAAATAATGAAAAAATGAGAGAGAAAGAAAGAAACCAACTGTGAC	12
Db	61	CTGAAATGSGCCAAAATAATGAAAAAATGAGAGAGAAAGAAAGAAACCAACTGTGAC	12
Qy	121	CACGTTTGCATGTTTCGCTATCCAAAATGGCTTGATAGGTGTCTATATGTGTGGGGAC	18
Db	121	CACGTTTGCATGTTTCGCTATCCAAAATGGCTTGATAGGTGTCTATATGTGTGGGGAC	18
Qy	181	AATGGCTGCATATCCATCCATGAGAGTGCACCTCCCTCTCATGATGCTGGTTTTGAAACAT	24
Db	181	AATGGCTGCATATCCATCCATGAGAGTGCACCTCCCTCTCATGATGCTGGTTTTGAAACAT	24
Qy	241	GACAGATAGCTTTCGAATGAGAGATTTCAAAAACAAACTTTTCCAGTTATATATTA	30
Db	241	GACAGATAGCTTTCGAATGAGAGATTTCAAAAACAAACTTTTCCAGTTATATATTA	30
Qy	301	TCAAGATTTTCGAAACATACACAATTTTCATCAACCATCTGAGAGAGAAATGACAC	36
Db	301	TCAAGATTTTCGAAACATACACAATTTTCATCAACCATCTGAGAGAGAAATGACAC	36
Qy	361	GTATGCCATATTTACAGTGGATGCGTGGCGGCGGCGGGGCGTCTACATCCAGT	42
Db	361	GTATGCCATATTTACAGTGGATGCGTGGCGGCGGCGGGGCGTCTACATCCAGT	42
Qy	421	TTCAATTCGGGCTCGGCACAGAGACAGATACATAAATATGAAAAACAATTTTTCAC	48
Db	421	TTCAATTCGGGCTCGGCACAGAGACAGATACATAAATATGAAAAACAATTTTTCAC	48
Qy	481	TGCTATCATCGACAGAGATTTGGCGGTTTACGTGATGATACGTTGGGAGCTTAACAC	54
Db	481	TGCTATCATCGACAGAGATTTGGCGGTTTACGTGATGATACGTTGGGAGCTTAACAC	54
Qy	541	CCGGCTACAGAGATGTTCCAAAATCAATGAAGAAATTTGGCCACAAAATTTGGAATGT	60
Db	541	CCGGCTACAGAGATGTTCCAAAATCAATGAAGAAATTTGGCCACAAAATTTGGAATGT	60
Qy	601	CTTTCACTCAATAGCAACAATTTTACCGGTTTATATGATGGGGTTTACAGTGTGGAA	66
Db	601	CTTTCACTCAATAGCAACAATTTTACCGGTTTATATGATGGGGTTTACAGTGTGGAA	66
Qy	661	GCATACCTTGATTTTGGCCATCAGCCGTTCTTGAGACTTTACGCCCATCTGGGC	72
Db	661	GCATACCTTGATTTTGGCCATCAGCCGTTCTTGAGACTTTACGCCCATCTGGGC	72
Qy	721	AAAGATACTATTTTATTTACTTATTCGATTAAGAAACCTTTGGCTATGCAAAAGCTGAGCAGT	78
Db	721	AAAGATACTATTTTATTTACTTATTCGATTAAGAAACCTTTGGCTATGCAAAAGCTGAGCAGT	78
Qy	781	AACTGAAGAAGTCTTACGAGCAATCAGACGTGATGTGCTTTGGAGCAAAAAGAAAGA	84
Db	781	AACTGAAGAAGTCTTACGAGCAATCAGACGTGATGTGCTTTGGAGCAAAAAGAAAGA	84
Qy	841	ACTTGAAGGTACAAACAAAATTTTAAAGAAAGTAAAGAAATTTGGATTAAGAAAGCTAT	90
Db	841	ACTTGAAGGTACAAACAAAATTTTAAAGAAAGTAAAGAAATTTGGATTAAGAAAGCTAT	90
Qy	901	CACGGCCAACTTTCTATTTGTTGGTCCGCTTTCTTATTTGATCTATGATCATATATGCTGGC	96
Db	901	CACGGCCAACTTTCTATTTGTTGGTCCGCTTTCTTATTTGATCTATGATCATATATGCTGGC	96
Qy	961	TTTCTGGATGGGACCCCTTGCTCTCTCCAGTGAATTTCTATTTGACAAAGTACTCAC	102
Db	961	TTTCTGGATGGGACCCCTTGCTCTCTCCAGTGAATTTCTATTTGACAAAGTACTCAC	102
Qy	1021	TGCTCTCTTTCTGATTAATTTGGGCTTTTGTATTTGAGACAGGCATCTCCACAGCATTTGA	108
Db	1021	TGCTCTCTTTCTGATTAATTTGGGCTTTTGTATTTGAGACAGGCATCTCCACAGCATTTGA	108

QY	1081	AGCATTTGCAAGCCGAAGAGGACACTTATGAAATCTTCAAGATAAATGACAAATAACC	1140
Db	1081	AGCATTTGCAAGCGCAAGAGAGACACTTATGAAATCTTCAAGATAAATGACAAATAACC	1140
QY	1141	AAGATTGACAGCTATTCCGAAGGTGCAATTAACCAAGATAATTAAGGAAATTTGGA	1200
Db	1141	AAGATTGACAGCTATTCCGAAGGTGCAATTAACCAAGATAATTAAGGAAATTTGGA	1200
QY	1201	ATTCAAAAATGTTCACTTGATTACCTCTTGAAAAGAGCTTTAAGATCTTAAAGGCTC	1260
Db	1201	ATTCAAAAATGTTCACTTGATTACCTCTTGAAAAGAGCTTTAAGATCTTAAAGGCTC	1260
QY	1261	CAACCTGAAGGTTCAGACGTGGGCAACACAGTGGCTGTTGGAAACAGTGGCTGGGGA	1320
Db	1261	CAACCTGAAGGTTCAGACGTGGGCAACACAGTGGCTGTTGGAAACAGTGGCTGGGGA	1320
QY	1321	GAGCAGACCGTGCAGCTGATGTCAGAGGCTGTATGACCCCAATGGCAATGGCTGTAT	1380
Db	1321	GAGCAGACCGTGCAGCTGATGTCAGAGGCTGTATGACCCCAATGGCAATGGCTGTAT	1380
QY	1381	TGATGACAGGACATTAGGACCAATAATGTAGGCATCTTGGGAAATTTACTGGTGTGT	1440
Db	1381	TGATGACAGGACATTAGGACCAATAATGTAGGCATCTTGGGAAATTTACTGGTGTGT	1440
QY	1441	GAGTCAGAGCCCTGTGTTGTTGGCCACAGATAGCTGGAAGAACTTGGCTATGGCCGCA	1500
Db	1441	GAGTCAGAGCCCTGTGTTGTTGGCCACAGATAGCTGGAAGAACTTGGCTATGGCCGCA	1500
QY	1501	AAATCTCACCATGGATGTGATTTGGAAGAGCTGTTAAGGAAGCCAAATGCTATGATTTAT	1560
Db	1501	AAATCTCACCATGGATGTGATTTGGAAGAGCTGTTAAGGAAGCCAAATGCTATGATTTAT	1560
QY	1561	CATGAACTACTATATAATTTGACTCTGTGTTGGAGAGAGGGGCCCACTAGTGG	1620
Db	1561	CATGAACTACTATATAATTTGACTCTGTGTTGGAGAGAGGGGCCCACTAGTGG	1620
QY	1621	TGAGCAGAAATAGGATTCGCCATTGCTGGGCCCTGTGTTGCGCAACCCCAAGATCTTCT	1680
Db	1621	TGAGCAGAAATAGGATTCGCCATTGCTGGGCCCTGTGTTGCGCAACCCCAAGATCTTCT	1680
QY	1681	GCTGATGAGGCAACGCTGCACTGTGACACTGAAAGTGAAGAGGTTCAGTGGCCCT	1740
Db	1681	GCTGATGAGGCAACGCTGCACTGTGACACTGAAAGTGAAGAGGTTCAGTGGCCCT	1740
QY	1741	GGATAAGCCCAAGAAAAGCCCGGACATCACTTGTGTGATAGCTATCTGTTGTCTACAGTTG	1800
Db	1741	GGATAAGCCCAAGAAAAGCCCGGACATCACTTGTGTGATAGCTATCTGTTGTCTACAGTTG	1800
QY	1801	TAATGCCGATGTCATGCTGGTTTTGATGATGAGATCACTTGTGAGAAAGGAATCATGA	1860
Db	1801	TAATGCCGATGTCATGCTGGTTTTGATGATGAGATCACTTGTGAGAAAGGAATCATGA	1860
QY	1861	TGAACTCATGAAAGAGAGAGGCGCATTTACTTCAAACTTGTCAATGCAAGCAAGAGAAA	1920
Db	1861	TGAACTCATGAAAGAGAGAGGCGCATTTACTTCAAACTTGTCAATGCAAGCAAGAGAAA	1920
QY	1921	TGAAAATGAGTTAGAAAATGCCACGSGGTATCCAAAAGTGAAGTATGCTCTTGGAAAT	1980
Db	1921	TGAAAATGAGTTAGAAAATGCCACGSGGTATCCAAAAGTGAAGTATGCTCTTGGAAAT	1980
QY	1981	GTCCTCAAAAGATTCAGGCTCCAGTTAAATAAAAAGAGATCAACTGCGAGGAGTTATA	2040
Db	1981	GTCCTCAAAAGATTCAGGCTCCAGTTAAATAAAAAGAGATCAACTGCGAGGAGTTATA	2040
QY	2041	TGCACCAACAAGCCCAAGCAGAAAGCTTGTGTACAAAAGAGAGCTTGAATGAGATGTAC	2100
Db	2041	TGCACCAACAAGCCCAAGCAGAAAGCTTGTGTACAAAAGAGAGCTTGAATGAGATGTAC	2100
QY	2101	TCCAGTTCTCTTGGAGGATTCGAAGCTGAAGCTCAACTGATGGCCCTATTTTTGTGT	2160
Db	2101	TCCAGTTCTCTTGGAGGATTCGAAGCTGAAGCTCAACTGATGGCCCTATTTTTGTGT	2160
QY	2161	TGGTATATTTTGTGCTATTTAAACGAGGCTTCGACCAACGACATTTTCAATTAATTTTC	2220

|||||
Db 2161 TGGTATATTTGGCTATATTAAGGAGGCGCTGCAACGACATTTTCATATATTTTC 2220
Oy 2221 AAGGATTAATAGGATCTTTACCCGAGATGAGATCTGAAACAAAGACAGATATGTA 2280
Db 2221 AAGGATTAATAGGATCTTTACCCGAGATGAGATCTGAAACAAAGACAGATATGTA 2280
Oy 2281 CATGTTTCTGATGTTTCTAGTCTTGGAAATTTCTTTATTTACATTTTCTCCA 2340
Db 2281 CATGTTTCTGATGTTTCTAGTCTTGGAAATTTCTTTATTTACATTTTCTCCA 2340
Oy 2341 GGGCTTCACATTTTGGCAAGCTGGGAGATCTCTCAAGGGCGTGCATGATGTTTT 2400
Db 2341 GGGCTTCACATTTTGGCAAGCTGGGAGATCTCTCAAGGGCGTGCATGATGTTTT 2400
Oy 2401 CAATCCATGCTGAGACAGATGTCAGTGTGATGATGACCTTAAACACCACTGGAGC 2460
Db 2401 CAATCCATGCTGAGACAGATGTCAGTGTGATGATGACCTTAAACACCACTGGAGC 2460
Oy 2461 ATTGACAAACGAGCTTGCCAAATGATGCGCTCAAGTTAAAGGGCGTATAGTTCCAGCT 2520
Db 2461 ATTGACAAACGAGCTTGCCAAATGATGCGCTCAAGTTAAAGGGCGTATAGTTCCAGCT 2520
Oy 2521 TGTCTGATATACCAAGATATAGCAATCTTGGACAGGCAATTTATATCTTAATCTA 2580
Db 2521 TGTCTGATATACCAAGATATAGCAATCTTGGACAGGCAATTTATATCTTAATCTA 2580
Oy 2581 TGGTTGGCAATTAACACTTTTACTCTTAAATTTGTAATTTGCAATTTGCAATGAGAGT 2640
Db 2581 TGGTTGGCAATTAACACTTTTACTCTTAAATTTGTAATTTGCAATTTGCAATGAGAGT 2640
Oy 2641 TGTGTAATTAAGAAATGTTGCTGAGCAAGCACTGAAAGATAGAAAGAGCTAGAGAGC 2700
Db 2641 TGTGTAATTAAGAAATGTTGCTGAGCAAGCACTGAAAGATAGAAAGAGCTAGAGAGC 2700
Oy 2701 TGGGAAGATTCCTACAGAAAGCCATCGAAATCTCCGACCTGTTCTTGGCTGGGA 2760
Db 2701 TGGGAAGATTCCTACAGAAAGCCATCGAAATCTCCGACCTGTTCTTGGCTGGGA 2760
Oy 2761 GCAGAAATTTGAATCATGTATGACAGAGTTTGCAGTATACCATACGAAACTCTTTGAG 2820
Db 2761 GCAGAAATTTGAATCATGTATGACAGAGTTTGCAGTATACCATACGAAACTCTTTGAG 2820
Oy 2821 GAAAGCACACATCTTGGGGTCTCATTTTCTATCACCCAGGCATGATGATTTTCTTA 2880
Db 2821 GAAAGCACACATCTTGGGGTCTCATTTTCTATCACCCAGGCATGATGATTTTCTTA 2880
Oy 2881 TGTGGCTGTTTCCGGTTGGTGTGCTACTGTGGCAATGAGTTGATGAACCTTTCAGGA 2940
Db 2881 TGTGGCTGTTTCCGGTTGGTGTGCTACTGTGGCAATGAGTTGATGAACCTTTCAGGA 2940
Oy 2941 TGTCTGCTGTTTCCGGTTGGTGTGCTACTGTGGCAATGAGTTGATGAACCTTTCAGGA 3000
Db 2941 TGTCTGCTGTTTCCGGTTGGTGTGCTACTGTGGCAATGAGTTGATGAACCTTTCAGGA 3000
Oy 3001 ATTTGCTCCGATATGCAAAAGCCAAAGTATCAGAGCCCACTCATCATGATCATGTA 3060
Db 3001 ATTTGCTCCGATATGCAAAAGCCAAAGTATCAGAGCCCACTCATCATGATCATGTA 3060
Oy 3061 AAAAAGCCCTCTGATTCAGACTACAGCCCTCAGGGCTCAAGCCAAATACCTTGGAGG 3120
Db 3061 AAAAAGCCCTCTGATTCAGACTACAGCCCTCAGGGCTCAAGCCAAATACCTTGGAGG 3120
Oy 3121 AAATGTGACATTTATAGAGTCTGTTCACATCTTCCACTCGACAGACATCCCGTGT 3180
Db 3121 AAATGTGACATTTATAGAGTCTGTTCACATCTTCCACTCGACAGACATCCCGTGT 3180
Oy 3181 CCAGGGGCTGAGCTCGAGGTGAAGAGGCGCAGAGCTGGCCCTGTAGGTGACAGTGG 3240
Db 3181 CCAGGGGCTGAGCTCGAGGTGAAGAGGCGCAGAGCTGGCCCTGTAGGTGACAGTGG 3240
Oy 3241 CTGTGGGAAGACAGTGTTCAGTCTCTAGAGCGCTTATGACCCCTTGGCTGTTTC 3300
|||||

Db 3241 CTGTGGGAAGAGCACAGTGTGTTACGTTCTAGAGCGCTTCTATGACCCCTTGGCTGTTTC 3300
Oy 3301 AGTGTATTTGATGGCAAGATTAAGCACTGATATGCTCCAGTGGCTCGAGACACCT 3360
Db 3301 AGTGTATTTGATGGCAAGATTAAGCACTGATATGCTCCAGTGGCTCGAGACACCT 3360
Oy 3361 GGGCATGCTGTTCTCAGAGCCCTCTGTTGATGAGTGGAGATTCGCGAAGCTTTCCTTA 3420
Db 3361 GGGCATGCTGTTCTCAGAGCCCTCTGTTGATGAGTGGAGATTCGCGAAGCTTTCCTTA 3420
Oy 3421 TGGAGCAACAGCCGCTGCTATCATGAGAGATTAATGAGGACCCAGAGAGGCGCA 3480
Db 3421 TGGAGCAACAGCCGCTGCTATCATGAGAGATTAATGAGGACCCAGAGAGGCGCA 3480
Oy 3481 CATACACACTTCATCAGACACTCCCTGAGAAATATACACACAGAGTAGAGACAAAG 3540
Db 3481 CATACACACTTCATCAGACACTCCCTGAGAAATATACACACAGAGTAGAGACAAAG 3540
Oy 3541 AACCGAGCTCTGGTGGCCAGAAACAGGCGATTGCAATGCTGCGCTCTTGTAGACA 3600
Db 3541 AACCGAGCTCTGGTGGCCAGAAACAGGCGATTGCAATGCTGCGCTCTTGTAGACA 3600
Oy 3601 GCCTCATATTTTCTTTTGGATGAGCTACATCAGCTCGATACAGAAAGTGAAGAAAGT 3660
Db 3601 GCCTCATATTTTCTTTTGGATGAGCTACATCAGCTCGATACAGAAAGTGAAGAAAGT 3660
Oy 3661 TGTCCAGAAAGCCCTGGACAAAGCCAGAGAGGCGCACCTCATGTTGATGCTCCACCG 3720
Db 3661 TGTCCAGAAAGCCCTGGACAAAGCCAGAGAGGCGCACCTCATGTTGATGCTCCACCG 3720
Oy 3721 CTGTGCAACATCCAGATGACATTTATAGTGTGTTTCAATGGCAAGTCAAGGA 3780
Db 3721 CTGTGCAACATCCAGATGACATTTATAGTGTGTTTCAATGGCAAGTCAAGGA 3780
Oy 3781 GCATGAGCAACATCAGACGCTGCGCCAGAAAGGCACTATTTTCCATGCTCAGT 3840
Db 3781 GCATGAGCAACATCAGACGCTGCGCCAGAAAGGCACTATTTTCCATGCTCAGT 3840
Oy 3841 CCAGGCTGAGCAAAAGCGCTAGTGAACCTGTGGCATATGACGCTTAAATTTTTTAAT 3900
Db 3841 CCAGGCTGAGCAAAAGCGCTAGTGAACCTGTGGCATATGACGCTTAAATTTTTTAAT 3900
Oy 3901 ATTTGTTTAAACATGSCATTTAATCAAGTTTAAAGGTGAGCACTTACTGAAAAACT 3960
Db 3901 ATTTGTTTAAACATGSCATTTAATCAAGTTTAAAGGTGAGCACTTACTGAAAAACT 3960
Oy 3961 ATGTGAACACTGTTTAAACATTTCTTGGCTGCAAGTGAAGATCATCCACCAAGTTCA 4020
Db 3961 ATGTGAACACTGTTTAAACATTTCTTGGCTGCAAGTGAAGATCATCCACCAAGTTCA 4020
Oy 4021 AGCTCTCAGATTTTAAATTAAGAACCAAGAAACATTAATCTGATGAAATTAATAC 4080
Db 4021 AGCTCTCAGATTTTAAATTAAGAACCAAGAAACATTAATCTGATGAAATTAATAC 4080
Oy 4081 TGTGTTAAATGCAATTAATTAATTAAGAGTAATCAAGTAGATTTGTTAATTAATTTG 4140
Db 4081 TGTGTTAAATGCAATTAATTAATTAAGAGTAATCAAGTAGATTTGTTAATTAATTTG 4140
Oy 4141 TATAATTTTGTATATTTTAAATTTTGAATTAAGTACTGCTTGTGCAAGATTAATGAG 4200
Db 4141 TATAATTTTGTATATTTTAAATTTTGAATTAAGTACTGCTTGTGCAAGATTAATGAG 4200
Oy 4201 GTAAAGTACTGATGTTTGAATTAAGTACTGATTAATTAATTAATTAATTAATTAATC 4260
Db 4201 GTAAAGTACTGATGTTTGAATTAAGTACTGATTAATTAATTAATTAATTAATTAATC 4260
Oy 4261 AAAAAAAAAAAAAAAAAA 4279
Db 4261 AAAAAAAAAAAAAAAAAA 4279
RESULT 4
AX105082

LOCUS	AX105082	4279 bp	DNA	linear	pat 30-APR-2001
DEFINITION	Sequence 26 from Patent WO0123540.				
ACCESSION	AX105082				
VERSION	AX105082.1 GI:13921232				
KEYWORDS	.				
SOURCE	dog.				
ORGANISM	Canis familiaris				
	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Carnivora; Fissipedia; Canidae; Canis.				
REFERENCE	1 (bases 1 to 4279)				
AUTHORS	Stocker,P.J., Steimel-Crespi,D.T., Crespi,C.L., Reif,T.C. and Patten,C.J.				
TITLE	P-glycoproteins and uses thereof				
JOURNAL	Patent: WO 0123540-A 26 05-APR-2001;				
FEATURES	GENESTR CORPORATION (US)				
source	location/Qualifiers				
	1..4279				

[illegible]

Query Match	99.9%	Score 4272.6	DB 6	Length 4279
Best Local Similarity	99.9%	Pred. No. 0		
Matches 4275	Conservative	0	Mismatches 4	Indels 0
				Gaps 0
QY	1	GGAGGCGAGGTGGGATGCATCTCTAAGAGGCGGTAAAGGGAGTCCAGAGAAGACTT	60	
Db	1	GGAGGCGAGGTGGGATGCATCTCTAAGAGGCGGTAAAGGGAGTCCAGAGAAGACTT	60	
QY	61	CTGGAATAATGGCCAAAAAAGTAAAAAAATAGAGAAAGAAAAAGAACCAACTGTGAC	120	
Db	61	CTGGAATAATGGCCAAAAAAGTAAAAAAATAGAGAAAGAAAAAGAACCAACTGTGAC	120	
QY	121	CACGTTGCAATGTTTGGCTATTCAAAATGGCTTGATAGTGGTATATGTTGGTGGGAC	180	
Db	121	CACGTTGCAATGTTTGGCTATTCAAAATGGCTTGATAGTGGTATATGTTGGTGGGAC	180	
QY	181	AATGGCTGCATCATCATGAGGCGACCTCCCTCTATGATGCTGTTTTGGAACAT	240	
Db	181	AATGGCTGCATCATCATGAGGCGACCTCCCTCTATGATGCTGTTTTGGAACAT	240	
QY	241	GACAGATACCTTTGCAAAATGCAGAAATTTAAGAAACAAAACCTTTCCAGTTATATTAA	300	
Db	241	GACAGATACCTTTGCAAAATGCAGAAATTTAAGAAACAAAACCTTTCCAGTTATATTAA	300	

QY	301	1GAAAGTATTAGCACAATATACAAACATTTTCATCAACCACTCTGAGAGAGAAATGACCAC	366
Db	301	TGAAGATTATTACGAACAAATATACAAATTTTCATCAACCACTCTGAGAGAGAAATGACCAC	366
QY	361	GTAAGCCATTTATTTTCAGTGGGATCGGCTGGCGTGCTGTGGCTGCTTACATCCAGT	420
Db	361	GTAATGCCATTTATTACAGTGGGATCGGCTGGCGTGCTGTGGCTGCTTACATCCAGT	420
QY	421	TTCAATTCGTGTGCTGGCAGCAGGAGAAACAGATACATCAAAATTTAAGAAACAATTTTTC	480
Db	421	TTCAATTCGTGTGCTGGCAGCAGGAGATTTGGCTGTGGCTGTGGCTGCTTACATCCAGT	480
QY	481	TGCTATCATGCGACAGGAGATTTGGCTGTGGCTGTGGCTGTGGCTGTGGCTGTGGCT	540
Db	481	TGCTATCATGCGACAGGAGATTTGGCTGTGGCTGTGGCTGTGGCTGTGGCTGTGGCT	540
QY	541	CCGGCTCAACAGCAGATGTCACAAAATCATGAAGAAATTTGGCACAATAATTTGAATGTT	600
Db	541	CCGGCTCAACAGCAGATGTCACAAAATCATGAAGAAATTTGGCACAATAATTTGAATGTT	600
QY	601	CTTTACATCAATACCAACATTTTTCACCGGTTTATAGTGGGTTTACAGTGTGGAA	660
Db	601	CTTTACATCAATACCAACATTTTTCACCGGTTTATAGTGGGTTTACAGTGTGGAA	660
QY	661	GCTAACCTTTGTATTTTGGCCATCGCCGCTTCTTGAGCTTTCAGCCGCATCGTGGGC	720
Db	661	GCTAACCTTTGTATTTTGGCCATCGCCGCTTCTTGAGCTTTCAGCCGCATCGTGGGC	720
QY	721	AAAGTACTATCTTCAATTACTGATTAAGAAACTCTTGCCCTATGCAAAAAGCTGAGAGCT	780
Db	721	AAAGTACTATCTTCAATTACTGATTAAGAAACTCTTGCCCTATGCAAAAAGCTGAGAGCT	780
QY	781	AGCTGAAGAGTCTGTGACGCAATGAGAACTGATTTGCTTTGAGGACAAAAGAAAGA	840
Db	781	AGCTGAAGAGTCTGTGACGCAATGAGAACTGATTTGCTTTGAGGACAAAAGAAAGA	840
QY	841	ACTTGAAGAGTACACAAAATTTAGAAAGAGCTAAAGAAATTTGGATTAAGAAAGCTAT	900
Db	841	ACTTGAAGAGTACACAAAATTTAGAAAGAGCTAAAGAAATTTGGATTAAGAAAGCTAT	900
QY	901	CACGGCCACATTTCTATTTGGGCGCGCTTCTATGATCTATGATCATATATGCTGCGC	960
Db	901	CACGGCCACATTTCTATTTGGGCGCGCTTCTATGATCTATGATCATATATGCTGCGC	960
QY	961	TTTTCTGTGTGGGACCTCTTGTGCTCTCCATGATTAATTTCTATTTGGACAAAGTACTCAC	1020
Db	961	TTTTCTGTGTGGGACCTCTTGTGCTCTCCATGATTAATTTCTATTTGGACAAAGTACTCAC	1020
QY	1021	TGCTCTCTTTCTGTATTAATTTGGGGCTTTTAGTATTTGACAGCCATCCCAAGCATTTGA	1080
Db	1021	TGCTCTCTTTCTGTATTAATTTGGGGCTTTTAGTATTTGACAGCCATCCCAAGCATTTGA	1080
QY	1081	AGCATTTGCCAAAGCAGAGAGAGACACTTATGAATTTCTCAAGATTAATTTGACAAATAAC	1140
Db	1081	AGCATTTGCCAAAGCAGAGAGAGACACTTATGAATTTCTCAAGATTAATTTGACAAATAAC	1140
QY	1141	AAGCATTTGACAGCTATTTGGAAGGTGAGATTAACCAAGATTAATTTAAGGAAATTTTGA	1200
Db	1141	AAGCATTTGACAGCTATTTGGAAGGTGAGATTAACCAAGATTAATTTAAGGAAATTTTGA	1200
QY	1201	ATTCAAAAATTTCTACCTAGTTACCTCTTCTGAAAAGATTAAGATCTTAAAGGCT	1260
Db	1201	ATTCAAAAATTTCTACCTAGTTACCTCTTCTGAAAAGATTAAGATCTTAAAGGCT	1260
QY	1261	CAACCTTAAGTTCAGAGTGGGCAACACAGTGGCTGTGTTGGAAACAGTGGCTGGGAA	1320
Db	1261	CAACCTTAAGTTCAGAGTGGGCAACACAGTGGCTGTGTTGGAAACAGTGGCTGGGAA	1320
QY	1321	GAGCAGACCGTGCAGCTGTATGACAGAGCTCTTAAGACCCCAAGATGGCATGTGCTAT	1380
Db	1321	GAGCAGACCGTGCAGCTGTATGACAGAGCTCTTAAGACCCCAAGATGGCATGTGCTAT	1380

1381 TGATGACAGACATTTAGACCAATTAATGTAAAGCATCTTCGGGAATTTACTGTGTGCT 1440
1381 TGATGACAGACATTTAGACCAATTAATGTAAAGCATCTTCGGGAATTTACTGTGTGCT 1440
1441 GAGTCAGAGAGCCTGTGTGTTTGGCCACACAGATAGCTGAAACAACTTCGATAGGCGCGCA 1500
1441 GAGTCAGAGAGCCTGTGTGTTTGGCCACACAGATAGCTGAAACAACTTCGATAGGCGCGCA 1500
1501 AAATGTCCACATGATGATGATTTGAAAGAGCTGTTAAAGAACCCAAATGCTATGATTTTAT 1560
1501 AAATGTCCACATGATGATGATTTGAAAGAGCTGTTAAAGAACCCAAATGCTATGATTTTAT 1560
1561 CATGAACACTACCTAATTAATTTGACACTCTGTGTTGAGAGAGAGGGCCCAAGCTGATG 1620
1561 CATGAACACTACCTAATTAATTTGACACTCTGTGTTGAGAGAGAGGGCCCAAGCTGATG 1620
1621 TGACAGAAACAGAGAAATGCGCATTCGTCGGGCCCTGGTTGCAACCCCAAGATTTCTT 1680
1621 TGACAGAAACAGAGAAATGCGCATTCGTCGGGCCCTGGTTGCAACCCCAAGATTTCTT 1680
1681 GCTGGATGAGGCAAGCTCAGCTCTGACACTGAAAGTGAAGCAGTGGTTGAGGTGGCCCT 1740
1681 GCTGGATGAGGCAAGCTCAGCTCTGACACTGAAAGTGAAGCAGTGGTTGAGGTGGCCCT 1740
1741 GGATTAAGCCAGAAAGAGCCGAGCACTGATGATGATGATGATGATGATGATGATGATG 1800
1741 GGATTAAGCCAGAAAGAGCCGAGCACTGATGATGATGATGATGATGATGATGATGATG 1800
1801 TAATGCCGATGTCATTCCTGTTTGTATGATGATGATGATGATGATGATGATGATGATG 1860
1801 TAATGCCGATGTCATTCCTGTTTGTATGATGATGATGATGATGATGATGATGATGATG 1860
1861 TGAATCTGAAAGAGAGGCAATTAATCAAACTTGTCAAAATGACAGACAAAGAGAAA 1920
1861 TGAATCTGAAAGAGAGGCAATTAATCAAACTTGTCAAAATGACAGACAAAGAGAAA 1920
1861 TGAATCTGAAAGAGAGGCAATTAATCAAACTTGTCAAAATGACAGACAAAGAGAAA 1920
1921 TGAATCTGAAAGAGAGGCAATTAATCAAACTTGTCAAAATGACAGACAAAGAGAAA 1920
1921 TGAATCTGAAAGAGAGGCAATTAATCAAACTTGTCAAAATGACAGACAAAGAGAAA 1920
1981 GTCTCCAAAGATTTAGAGGCTCCAGTTTAATTAAGAGATCAACCTGCGAGAGATATACA 2040
1981 GTCTCCAAAGATTTAGAGGCTCCAGTTTAATTAAGAGATCAACCTGCGAGAGATATACA 2040
1981 GTCTCCAAAGATTTAGAGGCTCCAGTTTAATTAAGAGATCAACCTGCGAGAGATATACA 2040
2041 TGACACAAAGGCAAG 2100
2041 TGACACAAAGGCAAG 2100
2101 TCCAGTTTCCTCTGAGAGATTTCTGAACCTGAACCTGAATGGCCTTATTTTGTGTG 2160
2101 TCCAGTTTCCTCTGAGAGATTTCTGAACCTGAACCTGAATGGCCTTATTTTGTGTG 2160
2161 TGGATTAATTTTGTCTAATTAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2220
2161 TGGATTAATTTTGTCTAATTAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2220
2161 TGGATTAATTTTGTCTAATTAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2220
2161 TGGATTAATTTTGTCTAATTAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2220
2221 AAGGATTAATGAGATCTTAACCGAGATGAGAGATGCTGAACCAAGAGAGAGATAGTAA 2280
2221 AAGGATTAATGAGATCTTAACCGAGATGAGAGATGCTGAACCAAGAGAGAGATAGTAA 2280
2281 CATGTTTCTGTATGTTTCTAGTCCCTGGAATTAATTTCTTTTATTAATTTTCTTCCA 2340
2281 CATGTTTCTGTATGTTTCTAGTCCCTGGAATTAATTTCTTTTATTAATTTTCTTCCA 2340
2341 GGGCTTTCACATTTGGCAAGCTGGGAGATCTCACTAAGAGGCGTTCATATGATGTTT 2400
2341 GGGCTTTCACATTTGGCAAGCTGGGAGATCTCACTAAGAGGCGTTCATATGATGTTT 2400
2401 CAGATCCATGCTGAGAGAGAGATGACAGTGTGATGACCTTAAACCAACCACTGAGAC 2460
2401 CAGATCCATGCTGAGAGAGAGATGACAGTGTGATGACCTTAAACCAACCACTGAGAC 2460
2461 ATTGACAAACAGAGCTTGGCAATGATGGGCTCAAGTTAAAGGGGCTATAGGTTCCAGGCT 2520

2461 ATTGACAAACAGAGCTTGGCAATGATGGGCTCAAGTTAAAGGGGCTATAGGTTCCAGGCT 2520
2521 TGCTGTTCATTACCCGAATATAGCAAAATCTTGGAGACAGCATTTATTAATCCTTATTTA 2580
2521 TGCTGTTCATTACCCGAATATAGCAAAATCTTGGAGACAGCATTTATTAATCCTTATTTA 2580
2581 TGGTGGCAATTAACACTTTTACTGTGACAAACACTGAAAGATTAAGAAAGAGCTAGAGAGC 2640
2581 TGGTGGCAATTAACACTTTTACTGTGACAAACACTGAAAGATTAAGAAAGAGCTAGAGAGC 2640
2641 TGTGGAATGAATATGTTGTCTGACAAAGCACTGAAAGATTAAGAAAGAGCTAGAGAGC 2700
2641 TGTGGAATGAATATGTTGTCTGACAAAGCACTGAAAGATTAAGAAAGAGCTAGAGAGC 2700
2701 TGGGAGATTTGCTACAGAAACCATGCAAACTTCGAACTGTGTTCTTTGACTCGGGA 2760
2701 TGGGAGATTTGCTACAGAAACCATGCAAACTTCGAACTGTGTTCTTTGACTCGGGA 2760
2761 GCAGAGTTTGAATATCATGTATGACAGAGATTGCAAGTACCATACAGAACTTTGAG 2820
2761 GCAGAGTTTGAATATCATGTATGACAGAGATTGCAAGTACCATACAGAACTTTGAG 2820
2821 GAAAGCACACATCTTGGGGTCTCATTTTCTATCAACCGAGCAATGATATTTTCTA 2880
2821 GAAAGCACACATCTTGGGGTCTCATTTTCTATCAACCGAGCAATGATATTTTCTA 2880
2881 TGCTGGCTGTTCCGGTTTGGTGGCTACTTGTGGCAANTGATTCATTAATTTTCAGA 2940
2881 TGCTGGCTGTTCCGGTTTGGTGGCTACTTGTGGCAANTGATTCATTAATTTTCAGA 2940
2941 TGCTGTTTGGTATTCAGCATTTGCTTGGGCAATGATTCAGTGGGCGAGTCACTTC 3000
2941 TGCTGTTTGGTATTCAGCATTTGCTTGGGCAATGATTCAGTGGGCGAGTCACTTC 3000
3001 ATTTGCTCCTGACTATGCAAAAGCAAGTATCAGAGCCAGCTCATCTGATATTA 3060
3001 ATTTGCTCCTGACTATGCAAAAGCAAGTATCAGAGCCAGCTCATCTGATATTA 3060
3061 AAAAGCCCTGTGATTTGACAGCTACAGCCCTCAGGCGCTCAAGCAATACCTTGAAG 3120
3061 AAAAGCCCTGTGATTTGACAGCTACAGCCCTCAGGCGCTCAAGCAATACCTTGAAG 3120
3121 AAATGTGACATTTAATGAGTGTGTTCAACTATCCACTGACAGACATCCCGTGTG 3180
3121 AAATGTGACATTTAATGAGTGTGTTCAACTATCCACTGACAGACATCCCGTGTG 3180
3181 CCAAGGCGTGAAGCTGAGGTGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 3240
3181 CCAAGGCGTGAAGCTGAGGTGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 3240
3241 CTGTGGGAAGAGCAAGTTTGTAGCTCTAGAGCGCTTCTATGACCCCTGGCGTGTG 3300
3241 CTGTGGGAAGAGCAAGTTTGTAGCTCTAGAGCGCTTCTATGACCCCTGGCGTGTG 3300
3301 AGTGTATATGATGGAAGAGATTAAGCACTGAAATGTCAGTGGCTCCAGACACACT 3360
3301 AGTGTATATGATGGAAGAGATTAAGCACTGAAATGTCAGTGGCTCCAGACACACT 3360
3361 GGGCATCTGTCTCAG 3420
3361 GGGCATCTGTCTCAG 3420
3421 TGGAGACAAAGCGCGGCTGTATCATGATGAAGATTAATGAGAGAGAGAGAGAGAGAG 3480
3421 TGGAGACAAAGCGCGGCTGTATCATGATGAAGATTAATGAGAGAGAGAGAGAGAGAG 3480
3481 CATACACCACTTCACTGAG 3540
3481 CATACACCACTTCACTGAG 3540
3541 AACCCAGCTCTGAGTGGAG 3600
3541 AACCCAGCTCTGAGTGGAG 3600

TTTCTCTGTGTGCGCAGAAACAGCGCATTTGCCATAGCTGGCGCTCTTGTGTAGACA	3600
TTTGTGCTTTGGATTAAGCTACATACAGCTCTGGATACAGAAAGTGA	3660
TTTGTGCTTTGGATTAAGCTACATACAGCTCTGGATTAAGAAAGTGA	3660
TTTGTGCTTTGGATTAAGCTACATACAGCTCTGGATTAAGAAAGTGA	3660
TTTGTGCTTTGGATTAAGCTACATACAGCTCTGGATTAAGAAAGTGA	3660
TTTGTGCTTTGGATTAAGCTACATACAGCTCTGGATTAAGAAAGTGA	3660
TTTGTGCTTTGGATTAAGCTACATACAGCTCTGGATTAAGAAAGTGA	3660
TTTGTGCTTTGGATTAAGCTACATACAGCTCTGGATTAAGAAAGTGA	3660
TTTGTGCTTTGGATTAAGCTACATACAGCTCTGGATTAAGAAAGTGA	3660
TTTGTGCTTTGGATTAAGCTACATACAGCTCTGGATTAAGAAAGTGA	3660
TTTGTGCTTTGGATTAAGCTACATACAGCTCTGGATTAAGAAAGTGA	3660
TTTGTGCTTTGGATTAAGCTACATACAGCTCTGGATTAAGAAAGTGA	3660
TTTGTGCTTTGGATTAAGCTACATACAGCTCTGGATTAAGAAAGTGA	3660
TTTGTGCTTTGGATTAAGCTACATACAGCTCTGGATTAAGAAAGTGA	3660
TTTGTGCTTTGGATTAAGCTACATACAGCTCTGGATTAAGAAAGTGA	3660
TTTGTGCTTTGGATTAAGCTACATACAGCTCTGGATTAAGAAAGTGA	3660
TTTGTGCTTTGGATTAAGCTACATACAGCTCTGGATTAAGAAAGTGA	3660
TTTGTGCTTTGGATTAAGCTACATACAGCTCTGGATTAAGAAAGTGA	3660
TTTGTGCTTTGGATTAAGCTACATACAGCTCTGGATTAAGAAAGTGA	3660
TTTGTGCTTTGGATTAAGCTACATACAGCTCTGGATTAAGAAAGTGA	3660
TTTGTGCTTTGGATTAAGCTACATACAGCTCTGGATTAAGAAAGTGA	3660
TTTGTGCTTTGGATTAAGCTACATACAGCTCTGGATTAAGAAAGTGA	3660
TTTGTGCTTTGGATTAAGCTACATACAGCTCTGGATTAAGAAAGTGA	3660
TTTGTGCTTTGGATTAAGCTACATACAGCTCTGGATTAAGAAAGTGA	3660
TTTGTGCTTTGGATTAAGCTACATACAGCTCTGGATTAAGAAAGTGA	3660
TTTGTGCTTTGGATTAAGCTACATACAGCTCTGGATTAAGAAAGTGA	3660
TTTGTGCTTTGGATTAAGCTACATACAGCTCTGGATTAAGAAAGTGA	3660
TTTGTGCTTTGGATTAAGCTACATACAGCTCTGGATTAAGAAAGTGA	3660
TTTGTGCTTTGGATTAAGCTACATACAGCTCTGGATTAAGAAAGTGA	3660
TTTGTGCTTTGGATTAAGCTACATACAGCTCTGGATTAAGAAAGTGA	3660
TTTGTGCTTTGGATTAAGCTACATACAGCTCTGGATTAAGAAAGTGA	3660
TTTGTGCTTTGGATTAAGCTACATACAGCTCTGGATTAAGAAAGTGA	3660
TTTGTGCTTTGGATTAAGCTACATACAGCTCTGGATTAAGAAAGTGA	3660
TTTGTGCTTTGGATTAAGCTACATACAGCTCTGGATTAAGAAAGTGA	3660
TTTGTGCTTTGGATTAAGCTACATACAGCTCTGGATTAAGAAAGTGA	3660
TTTGTGCTTTGGATTAAGCTACATACAGCTCTGGATTAAGAAAGTGA	3660
TTTGTGCTTTGGATTAAGCTACATACAGCTCTGGATTAAGAAAGTGA	3660
TTTGTGCTTTGGATTAAGCTACATACAGCTCTGGATTAAGAAAGTGA	3660
TTTGTGCTTTGGATTAAGCTACATACAGCTCTGGATTAAGAAAGTGA	3660
TTTGTGCTTTGGATTAAGCTACATACAGCTCTGGATTAAGAAAGTGA	3660
TTTGTGCTTTGGATTAAGCTACATACAGCTCTGGATTAAGAAAGTGA	3660
TTTGTGCTTTGGATTAAGCTACATACAGCTCTGGATTAAGAAAGTGA	3660
TTTGTGCTTTGGATTAAGCTACATACAGCTCTGGATTAAGAAAGTGA	3660
TTTGTGCTTTGGATTAAGCTACATACAGCTCTGGATTAAGAAAGTGA	3660
TTTGTGCTTTGGATTAAGCTACATACAGCTCTGGATTAAGAAAGTGA	3660
TTTGTGCTTTGGATTAAGCTACATACAGCTCTGGATTAAGAAAGTGA	3660
TTTGTGCTTTGGATTAAGCTACATACAGCTCTGGATTAAGAAAGTGA	3660
TTTGTGCTTTGGATTAAGCTACATACAGCTCTGGATTAAGAAAGTGA	3660
TTTGTGCTTTGGATTAAGCTACATACAGCTCTGGATTAAGAAAGTGA	3660
TTTGTGCTTTGGATTAAGCTACATACAGCTCTGGATTAAGAAAGTGA	3660
TTTGTGCTTTGGATTAAGCTACATACAGCTCTGGATTAAGAAAGTGA	3660
TTTGTGCTTTGGATTAAGCTACATACAGCTCTGGATTAAGAAAGTGA	3660
TTTGTGCTTTGGATTAAGCTACATACAGCTCTGGATTAAGAAAGTGA	3660
TTTGTGCTTTGGATTAAGCTACATACAGCTCTGGATTAAGAAAGTGA	3660
TTTGTGCTTTGGATTAAGCTACATACAGCTCTGGATTAAGAAAGTGA	3660
TTTGTGCTTTGGATTAAGCTACATACAGCTCTGGATTAAGAAAGTGA	3660
TTTGTGCTTTGGATTAAGCTACATACAGCTCTGGATTAAGAAAGTGA	3660
TTTGTGCTTTGGATTAAGCTACATACAGCTCTGGATTAAGAAAGTGA	3660
TTTGTGCTTTGGATTAAGCTACATACAGCTCTGGATTAAGAAAGTGA	3660
TTTGTGCTTTGGATTAAGCTACATACAGCTCTGGATTAAGAAAGTGA	3660
TTTGTGCTTTGGATTAAGCTACATACAGCTCTGGATTAAGAAAGTGA	3660
TTTGTGCTTTGGATTAAGCTACATACAGCTCTGGATTAAGAAAGTGA	3660
TTTGTGCTTTGGATTAAGCTACATACAGCTCTGGATTAAGAAAGTGA	3660
TTTGTGCTTTGGATTAAGCTACATACAGCTCTGGATTAAGAAAGTGA	3660
TTTGTGCTTTGGATTAAGCTACATACAGCTCTGGATTAAGAAAGTGA	3660
TTTGTGCTTTGGATTAAGCTACATACAGCTCTGGATTAAGAAAGTGA	3660
TTTGTGCTTTGGATTAAGCTACATACAGCTCTGGATTAAGAAAGTGA	3660
TTTGTGCTTTGGATTAAGCTACATACAGCTCTGGATTAAGAAAGTGA	3660
TTTGTGCTTTGGATTAAGCTACATACAGCTCTGGATTAAGAAAGTGA	3660
TTTGTGCTTTGGATTAAGCTACATACAGCTCTGGATTAAGAAAGTGA	3660
TTTGTGCTTTGGATTAAGCTACATACAGCTCTGGATTAAGAAAGTGA	3660
TTTGTGCTTTGGATTAAGCTACATACAGCTCTGGATTAAGAAAGTGA	3660
TTTGTGCTTTGGATTAAGCTACATACAGCTCTGGATTAAGAAAGTGA	3660
TTTGTGCTTTGGATTAAGCTACATACAGCTCTGGATTAAGAAAGTGA	3660
TTTGTGCTTTGGATTAAGCTACATACAGCTCTGGATTAAGAAAGTGA	3660
TTTGTGCTTTGGATTAAGCTACATACAGCTCTGGATTAAGAAAGTGA	3660
TTTGTGCTTTGGATTAAGCTACATACAGCTCTGGATTAAGAAAGTGA	3660
TTTGTGCTTTGGATTAAGCTACATACAGCTCTGGATTAAGAAAGTGA	3660
TTTGTGCTTTGGATTAAGCTACATACAGCTCTGGATTAAGAAAGTGA	3660
TTTGTGCTTTGGATTAAGCTACATACAGCTCTGGATTAAGAAAGTGA	3660
TTTGTGCTTTGGATTAAGCTACATACAGCTCTGGATTAAGAAAGTGA	3660
TTTGTGCTTTGGATTAAGCTACATACAGCTCTGGATTAAGAAAGTGA	3660
TTTGTGCTTTGGATTAAGCTACATACAGCTCTGGATTAAGAAAGTGA	3660
TTTGTGCTTTGGATTAAGCTACATACAGCTCTGGATTAAGAAAGTGA	3660
TTTGTGCTTTGGATTAAGCTACATACAGCTCTGGATTAAGAAAGTGA	3660
TTTGTGCTTTGGATTAAGCTACATACAGCTCTGGATTAAGAAAGTGA	3660
TTTGTGCTTTGGATTAAGCTACATACAGCTCTGGATTAAGAAAGTGA	3660
TTTGTGCTTTGGATTAAGCTACATACAGCTCTGGATTAAGAAAGTGA	3660
TTTGTGCTTTGGATTAAGCTACATACAGCTCTGGATTAAGAAAGTGA	3660
TTTGTGCTTTGGATTAAGCTACATACAGCTCTGGATTAAGAAAGTGA	3660
TTTGTGCTTTGGATTAAGCTACATACAGCTCTGGATTAAGAAAGTGA	3660
TTTGTGCTTTGGATTAAGCTACATACAGCTCTGGATTAAGAAAGTGA	3660
TTTGTGCTTTGGATTAAGCTACATACAGCTCTGGATTAAGAAAGTGA	3660
TTTGTGCTTTGGATTAAGCTACATACAGCTCTGGATTAAGAAAGTGA	3660
TTTGTGCTTTGGATTAAGCTACATACAGCTCTGGATTAAGAAAGTGA	3660
TTTGTGCTTTGGATTAAGCTACATACAGCTCTGGATTAAGAAAGTGA	3660
TTTGTGCTTTGGATTAAGCTACATACAGCTCTGGATTAAGAAAGTGA	3660
TTTGTGCTTTGGATTAAGCTACATACAGCTCTGGATTAAGAAAGTGA	3660
TTTGTGCTTTGGATTAAGCTACATACAGCTCTGGATTAAGAAAGTGA	3660
TTTGTGCTTTGGATTAAGCTACATACAGCTCTGGATTAAGAAAGTGA	3660
TTTGTGCTTTGGATTAAGCTACATACAGCTCTGGATTAAGAAAGTGA	3660
TTTGTGCTTTGGATTAAGCTACATACAGCTCTGGATTAAGAAAGTGA	3660
TTTGTGCTTTGGATTAAGCTACATACAGCTCTGGATTAAGAAAGTGA	3660
TTTGTGCTTTGGATTAAGCTACATACAGCTCTGGATTAAGAAAGTGA	3660
TTTGTGCTTTGGATTAAGCTACATACAGCTCTGGATTAAGAAAGTGA	3660
TTTGTGCTTTGGATTAAGCTACATACAGCTCTGGATTAAGAAAGTGA	3660
TTTGTGCTTTGGATTAAGCTACATACAGCTCTGGATTAAGAAAGTGA	3660
TTTGTGCTTTGGATTAAGCTACATACAGCTCTGGATTAAGAAAGTGA	3660
TTTGTGCTTTGGATTAAGCTACATACAGCTCTGGATTAAGAAAGTGA	3660
TTTGTGCTTTGGATTAAGCTACATACAGCTCTGGATTAAGAAAGTGA	3660
TTTGTGCTTTGGATTAAGCTACATACAGCTCTGGATTAAGAAAGTGA	3660
TTTGTGCTTTGGATTAAGCTACATACAGCTCTGGATTAAGAAAGTGA	3660
TTTGTGCTTTGGATTAAGCTACATACAGCTCTGGATTAAGAAAGTGA	3660
TTTGTGCTTTGGATTAAGCTACATACAGCTCTGGATTAAGAAAGTGA	3660
TTTGTGCTTTGGATTAAGCTACATACAGCTCTGGATTAAGAAAGTGA	3660
TTTGTGCTTTGGATTAAGCTACATACAGCTCTGGATTAAGAAAGTGA	3660
TTTGTGCTTTGGATTAAGCTACATACAGCTCTGGATTAAGAAAGTGA	3660
TTTGTGCTTTGGATTAAGCTACATACAGCTCTGGATTAAGAAAGTGA	3660
TTTGTGCTTTGGATTAAGCTACATACAGCTCTGGATTAAGAAAGTGA	3660
TTTGTGCTTTGGATTAAGCTACATACAGCTCTGGATTAAGAAAGTGA	3660
TTTGTGCTTTGGATTAAGCTACATACAGCTCTGGATTAAGAAAGTGA	3660
TTTGTGCTTTGGATTAAGCTACATACAGCTCTGGATTAAGAAAGTGA	3660
TTTGTGCTTTGGATTAAGCTACATACAGCTCTGGATTAAGAAAGTGA	3660
TTTGTGCTTTGGATTAAGCTACATACAGCTCTGGATTAAGAAAGTGA	3660
TTTGTGCTTTGGATTAAGCTACATACAGCTCTGGATTAAGAAAGTGA	3660
TTTGTGCTTTGGATTAAGCTACATACAGCTCTGGATTAAGAAAGTGA	3660
TTTGTGCTTTGGATTAAGCTACATACAGCTCTGGATTAAGAAAGTGA	3660
TTTGTGCTTTGGATTAAGCTACATACAGCTCTGGATTAAGAAAGTGA	3660
TTTGTGCTTTGGATTAAGCTACATACAGCTCTGGATTAAGAAAGTGA	3660
TTTGTGCTTTGGATTAAGCTACATACAGCTCTGGATTAAGAAAGTGA	3660
TTTGTGCTTTGGATTAAGCTACATACAGCTCTGGATTAAGAAAGTGA	3660
TTTGTGCTTTGGATTAAGCTACATACAGCTCTGGATTAAGAAAGTGA	3660
TTTGTGCTTTGGATTAAGCTACATACAGCTCTGGATTAAGAAAGTGA	3660
TTTGTGCTTTGGATTAAGCTACATACAGCTCTGGATTAAGAAAGTGA	3660
TTTGTGCTTTGGATTAAGCTACATACAGCTCTGGATTAAGAAAGTGA	3660
TTTGTGCTTTGGATTAAGCTACATACAGCTCTGGATTAAGAAAGTGA	3660
TTTGTGCTTTGGATTAAGCTACATACAGCTCTGGATTAAGAAAGTGA	3660
TTTGTGCTTTGGATTAAGCTACATACAGCTCTGGATTAAGAAAGTGA	3660
TTTGTGCTTTGGATTAAGCTACATACAGCTCTGGATTAAGAAAGTGA	3660
TTTGTGCTTTGGATTAAGCTACATACAGCTCTGGATTAAGAAAGTGA	3660
TTTGTGCTTTGGATTAAGCTACATACAGCTCTGGATTAAGAAAGTGA	3660
TTTGTGCTTTGGATTAAGCTACATACAGCTCTGGATTAAGAAAGTGA	3660
TTTGTGCTTTGGATTAAGCTACATACAGCTCTGGATTAAGAAAGTGA	3660
TTTGTGCTTTGGATTAAGCTACATACAGCTCTGGATTAAGAAAGTGA	3660
TTTGTGCTTTGGATTAAGCTACATACAGCTCTGGATTAAGAAAGTGA	3660
TTTGTGCTTTGGATTAAGCTACATACAGCTCTGGATTAAGAAAGTGA	3660
TTTGTGCTTTGGATTAAGCTACATACAGCTCTGGATTAAGAAAGTGA	3660
TTTGTGCTTTGGATTAAGCTACATACAGCTCTGGATTAAGAAAGTGA	3660
TTTGTGCTTTGGATTAAGCTACATACAGCTCTGGATTAAGAAAGTGA	3660
TTTGTGCTTTGGATTAAGCTACATACAGCTCTGGATTAAGAAAGTGA	3660
TTTGTGCTTTGGATTAAGCTACATACAGCTCTGGATTAAGAAAGTGA	3660
TTTGTGCTTTGGATTAAGCTACATACAGCTCTGGATTAAGAAAGTGA	3660
TTTGTGCTTTGGATTAAGCTACATACAGCTCTGGATTAAGAAAGTGA	3660
TTTGTGCTTTGGATTAAGCTACATACAGCTCTGGATTAAGAAAGTGA	3660
TTTGTGCTTTGGATTAAGCTACATACAGCTCTGGATTAAGAAAGTGA	3660
TTTGTGCTTTGGATTAAGCTACATACAGCTCTGGATTAAGAAAGTGA	3660
TTTGTGCTTTGGATTAAGCTACATACAGCTCTGGATTAAGAAAGTGA	3660
TTTGTGCTTTGGATTAAGCTACATACAGCTCTGGATTAAGAAAGTGA	3660
TTTGTGCTTTGGATTAAGCTACATACAGCTCTGGATTAAGAAAGTGA	3660
TTTGTGCTTTGGATTAAGCTACATACAGCTCTGGATTAAGAAAGTGA	3660
TTTGTGCTTTGGATTAAGCTACATACAGCTCTGGATTAAGAAAGTGA	3660
TT	

[illegible]

Db 533 CTATCATGCGACAGAGATTGGCTGGTTGACGTCGATGACCTTGGGAGGCTTAACAACC 592
Qy 543 GGCTCAGACGATGTCCTCCAAATCAATGAGAAATTGGCGACAAATTTGGAATGTTCT 602
Db 593 GGCTCAGACGATGTCCTCCAAATCAATGAGAAATTGGCGACAAATTTGGAATGTTCT 652
Qy 603 TTGACTCAATAGCAACATTTTTCACCGGTTTATAGTGGGGTTTACAGTGGTTGGAAGC 662
Db 653 TTCAATCAATAGCAACATTTTTCACCGGTTTATAGTGGGGTTTACAGTGGTTGGAAGC 712
Qy 663 TAACCCCTGGATTGGGCCATCAGCCCTGTTCTTGGACTTTACGCCCATCTGGGCA 722
Db 713 TAACCCCTGGATTGGGCCATCAGCCCTGTTCTTGGACTTTACGCCCATCTGGGCA 772
Qy 723 AGATCTATCTCTCATTTACTGATPAAGAACTCTTGGCCATATGCAAAAGGTGAGCAATG 782
Db 773 AGATCTATCTCTCATTTACTGATPAAGAACTCTTGGCCATATGCAAAAGGTGAGCAATG 832
Qy 783 CTGAAGAAGTCTTAGCAGCAATCAAGACCTGATGCTTGGAGGACAAAGAAAGAAC 842
Db 833 CTGAAGAAGTCTTAGCAGCAATCAAGACCTGATGCTTGGAGGACAAAGAAAGAAC 892
Qy 843 TTGAAGGTACACAAAAATTTAGAGAGCTAAAGGAATTGGGATPAAGAAAGCTATCA 902
Db 893 TTGAAGGTACACAAAAATTTAGAGAGCTAAAGGAATTGGGATPAAGAAAGCTATCA 952
Qy 903 CGGCCACATTTCTATTGGTCCCTCTTCTTATGATCTATGATCAATGCTATGCTGCTT 962
Db 953 CGGCCACATTTCTATTGGTCCCTCTTCTTATGATCTATGATCAATGCTATGCTGCTT 1012
Qy 963 TCTGGTATGGGACCTCTTGGTCCCTCTTCTTCAAGTAATTTCTATTGGAAGTACTACTG 1022
Db 1013 TCTGGTATGGGACCTCTTGGTCCCTCTTCTTCAAGTAATTTCTATTGGAAGTACTACTG 1072
Qy 1023 TCTTCTTCTGTAATTAATTGGGGCTTTTACTATTGGACAGAGCTTCCCAAGCTTAAG 1082
Db 1073 TCTTCTTCTGTAATTAATTGGGGCTTTTACTATTGGAGAGGCTATCCCAAGCTTAAG 1132
Qy 1083 CATTTCAAAAGCAGAGAGAGAGCTTATGAAATCTTCAAGATTAATTGACATTAACCA 1142
Db 1133 CATTTCAAAAGCAGAGAGAGAGCTTATGAAATCTTCAAGATTAATTGACATTAACCA 1192
Qy 1143 GCATTGACAGCTATTTGGAAGATGAGACATTAACAGATTAATTGAGGAATTTGCAAT 1202
Db 1193 GCATTGACAGCTATTTGGAAGATGAGACATTAACAGATTAATTGAGGAATTTGCAAT 1252
Qy 1203 TCAAAAATGTTCACTTCACTTACCCTTCTCGAAAAAGATTAAAGTCTTAAAGGCTCA 1262
Db 1253 TCAAAAATGTTCACTTCACTTACCCTTCTCGAAAAAGATTAAAGTCTTAAAGGCTCA 1312
Qy 1263 ACCTGAAGTTCAAGAGTGGGACAGACAGTGGCGCTGGTTGGGAACAGTGGCTGGGAGA 1322
Db 1313 ACCTGAAGTTCAAGAGTGGGACAGACAGTGGCGCTGGTTGGGAACAGTGGCTGGGAGA 1372
Qy 1323 GCAGAGACCGGTGAGCTGATGCAGAGGCTATGACCCACAGATGGCATGCTGTATTTG 1382
Db 1373 GCAGAGACCGGTGAGCTGATGCAGAGGCTATGACCCACAGATGGCATGCTGTATTTG 1432
Qy 1383 ATGAGACAGACATTAGGACCATTAATGTAAGCATTTTGGGAAATTAAGTGGTGGTGA 1442
Db 1433 ATGAGACAGACATTAGGACCATTAATGTAAGCATTTTGGGAAATTAAGTGGTGGTGA 1492
Qy 1443 GTTCAGAGCCTGTGTGTTGGCCACCGAGATGTAAGAAATTCCTATAGCCCGGGA 1502
Db 1493 GTTCAGAGCCTGTGTGTTGGCCACCGAGATGTAAGAAATTCCTATAGCCCGGGA 1552
Qy 1503 ATGTCACCATGATGAGATTGAGAAAGCTGTTAAGGAAGCAATGCCATATGATTTATCA 1562
Db 1553 ATGTCACCATGATGAGATTGAGAAAGCTGTTAAGGAAGCAATGCCATATGATTTATCA 1612
Qy 1563 TGAACCTACCTAATTAATTTGACACTGTGGTTGGAGAGAGAGGGCCCGGCTAGTGGTG 1622
Db 1613 TGAACCTACCTAATTAATTTGACACTGTGGTTGGAGAGAGAGGGCCCGGCTAGTGGTG 1672

Qy 1623 GACAGAAACAGAGAAATTCGCCATTTGCTCGGGCCCTGGTTCCGACCCCAAGATTTCTTCG 1682
Db 1673 GACAGAAACAGAGAAATTCGCCATTTGCTCGGGCCCTGGTTCCGACCCCAAGATTTCTTCG 1732
Qy 1683 TGGATGAGCAACGTCAGCTGACACTGAAAGTGAAGCAGTGGTTCAGTGGCCCTGG 1742
Db 1733 TGGATGAGCAACGTCAGCTGACACTGAAAGTGAAGCAGTGGTTCAGTGGCCCTGG 1792
Qy 1743 ATAGGCCAGAAAAAGCCGAGCTACCATTTGTATAGTCATCTGTTCTACAGTTGCTA 1802
Db 1793 ATAGGCCAGAAAAAGCCGAGCTACCATTTGTATAGTCATCTGTTCTACAGTTGCTA 1852
Qy 1803 ATGCCCATGTCATTTGCTGGTTTGGATGATGAGTCAATTCGAGAAAGAAATCATGATG 1862
Db 1853 ATGCCCATGTCATTTGCTGGTTTGGATGATGAGTCAATTCGAGAAAGAAATCATGATG 1912
Qy 1863 AACCTATGAAGAGAGAGGGCATTTACTTCAAACTTGTCCAAATGCAAGCAAGAGAAATG 1922
Db 1913 AACCTATGAAGAGAGAGGGCATTTACTTCAAACTTGTCCAAATGCAAGCAAGAGAAATG 1972
Qy 1923 AAATGAGTTAGAAAATGCCACGTGGGAATCCAAAGTGAAGTCAATGCTTGAATGT 1982
Db 1973 AAATGAGTTAGAAAATGCCACGTGGGAATCCAAAGTGAAGTCAATGCTTGAATGT 2032
Qy 1983 CTCCAAAAGATTCAAGGCTCCAGTTTAAATAAAAGAAAGATCACTGCGAGATATACATG 2042
Db 2033 CTCCAAAAGATTCAAGGCTCCAGTTTAAATAAAAGAAAGATCACTGCGAGATATACATG 2092
Qy 2043 CACCACAGGCCAAGACAGAAAGCTTGTACAAAAGAGGACTTGAATGAAATGTACTTC 2102
Db 2093 CACCACAGGCCAAGACAGAAAGCTTGTACAAAAGAGGACTTGAATGAAATGTACTTC 2152
Qy 2103 CAGTTTCCCTTCCGAGATCTGAAGCTGAATCACTGAATGAGCCTTATTTGTGCTTG 2162
Db 2153 CAGTTTCCCTTCCGAGATCTGAAGCTGAATCACTGAATGAGCCTTATTTGTGCTTG 2212
Qy 2163 GTATATTTTGTCTATTTAATAACGAGGCTTCGAACAGCAATTTCAATATATTTTCA 2222
Db 2213 GTATATTTTGTCTATTTAATAACGAGGCTTCGAACAGCAATTTCAATATATTTTCA 2272
Qy 2223 GATATATGAGGATCTTTACCGAGATGAGATCCCTGAAAACAAAACGAGAAATGATACA 2282
Db 2273 GATATATGAGGATCTTTACCGAGATGAGATCCCTGAAAACAAAACGAGAAATGATACA 2332
Qy 2283 TGTTTTCTGTAATGTTTCTAGTCTTGGAAATTAATTTCTTATTAACATTTTCTCCAG 2342
Db 2333 TGTTTTCTGTAATGTTTCTAGTCTTGGAAATTAATTTCTTATTAACATTTTCTCCAG 2392
Qy 2343 GCTTCACTTTGGCAAGGCTGGGAGATCCCTACCTAACCGGCTTCGATACATGCTTTCA 2402
Db 2393 GCTTCACTTTGGCAAGGCTGGGAGATCCCTACCTAACCGGCTTCGATACATGCTTTCA 2452
Qy 2403 GATCCATGCTGAGACAGAGATGTCAGTGTGTTGATGACCTTAAACACGACCTGAGCAT 2462
Db 2453 GATCCATGCTGAGACAGAGATGTCAGTGTGTTGATGACCTTAAACACGACCTGAGCAT 2512
Qy 2463 TGACAAACAGGCTTGGCAATGATGAGGCTCAAGTTAAAGGGGCTTATGCTTCAGGCTTG 2522
Db 2513 TGACAAACAGGCTTGGCAATGATGAGGCTCAAGTTAAAGGGGCTTATGCTTCAGGCTTG 2572
Qy 2523 CTGTCATTACCCGAATATAGCAAACTTGGGACAGGCTATTTATATCTTAAATCTATG 2582
Db 2573 CTGTCATTACCCGAATATAGCAAACTTGGGACAGGCTATTTATATCTTAAATCTATG 2632
Qy 2583 GTTGGCAATTAACCTTTTACTCTTAGCAATTTTACCATCTATTCGAATAGCAGAGTTG 2642
Db 2633 GTTGGCAATTAACCTTTTACTCTTAGCAATTTTACCATCTATTCGAATAGCAGAGTTG 2692
Qy 2643 TTGAATAAGAAATGTTGTCGACAAAGCACTGAAAGATTAAGAAAGAGTGAAGAGCTG 2702
Db 2693 TTGAATAAGAAATGTTGTCGACAAAGCACTGAAAGATTAAGAAAGAGTGAAGAGCTG 2752

[illegible]

QY	1923	AAATGAGTTAGAAAATGCCACTGSGTGAATCCAAAAGTAAAGTAACTGATGCTTTGGAAATGT	1982
Db	1973	AAATGAGTTAGAAAATGCCACTGSGTGAATCCAAAAGTAAAGTAACTGATGCTTTGGAAATGT	2032
QY	1983	CTCCAAAAGANTCAGGGTCCAGTTTAAATAAAAAGAGATCAACTCGCAGAGTATACATG	2042
Db	2033	CTCCAAAAGANTCAGGGTCCAGTTTAAATAAAAAGAGATCAACTCGCAGAGTATACATG	2092
QY	2043	CACCACAAGGCCAAGACAGAAACCTTGATACAAAAGAGAGACTGATAGTAGNTACCTC	2102
Db	2093	CACCACAAGGCCAAGACAGAAACCTTGATACAAAAGAGAGACTGATAGTAGNTACCTC	2152
QY	2103	CAGTTTCCTCTGAGAGATTTGAAAGCTGAACTCAACTGAAATGAGCTTAATTTGGTGTG	2162
Db	2153	CAGTTTCCTCTGAGAGATTTGAAAGCTGAACTCAACTGAAATGAGCTTAATTTGGTGTG	2212
QY	2163	GTAATATTTTGCTGATTAATTAACGGAGGCTCGCAACCGACTTTGCAATAAATTTTCA	2222
Db	2213	GTAATATTTTGCTGATTAATTAACGGAGGCTCGCAACCGACTTTGCAATAAATTTTCA	2272
QY	2223	GGATTAATAGGATCTTTAACCCGAGATGAGATTCCTGAAACAAAACGACAGAAATGTAACA	2282
Db	2273	GGATTAATAGGATCTTTAACCCGAGATGAGATTCCTGAAACAAAACGACAGAAATGTAACA	2332
QY	2283	TGTTTTCTGATTTGTTTCTAGTCCTGGAATATTTCTTTTAATTAATTTTCTCCAG	2342
Db	2333	TGTTTTCTGATTTGTTTCTAGTCCTGGAATATTTCTTTTAATTAATTTTCTCCAG	2392
QY	2343	GCTTACACTTTGGCAAAAGCTGGGAGATTCCTCACTAAGGGGCTTGATACATAGTGTTC	2402
Db	2393	GCTTACACTTTGGCAAAAGCTGGGAGATTCCTCACTAAGGGGCTTGATACATAGTGTTC	2452
QY	2403	GATCCATCTGAGACAGAGATGTGAGCGGTGATATACCTTAATAAACACCACTGAGAGAT	2462
Db	2453	GATCCATCTGAGACAGAGATGTGAGCGGTGATATACCTTAATAAACACCACTGAGAGAT	2512
QY	2463	TGACAACCCAGGCTTGCCAAATGATGCGGCTCAAGTTAAAGGGGCTTATAGTTCACAGGCTG	2522
Db	2513	TGACAACCCAGGCTTGCCAAATGATGCGGCTCAAGTTAAAGGGGCTTATAGTTCACAGGCTG	2572
QY	2523	CTGTCAATTACCCAGAAATATATAGCAAACTTTGGGACAGGCAATTAATATCTTAATCTATG	2582
Db	2573	CTGTCAATTACCCAGAAATATATAGCAAACTTTGGGACAGGCAATTAATATCTTAATCTATG	2632
QY	2583	GTGGGCAATTAACACTTTTACTCTTAAGCAATTGTACCATATTTGCAATATGCAAGAGTTG	2642
Db	2633	GTGGGCAATTAACACTTTTACTCTTAAGCAATTGTACCATATTTGCAATATGCAAGAGTTG	2692
QY	2643	TTGAATAAGAAAATGTTGTGTGGACAAGCACTGAAAGATTAAGAAAGAGACTGAAAGAGCTG	2702
Db	2693	TTGAATAAGAAAATGTTGTGTGGACAAGCACTGAAAGATTAAGAAAGAGACTGAAAGAGCTG	2752
QY	2703	GGAAAGTTCTACAGAAGCATCGAAACCTCGCAACGTTGTTCTTTTACTCGGAGAC	2762
Db	2753	GGAAAGTTCTACAGAAGCATCGAAACCTCGCAACGTTGTTCTTTTACTCGGAGAC	2812
QY	2763	AGAAAGTTAAATACATGATACAGAGAGTTTGCAAGTACATACAGAAATCTTTTGAAGA	2822
Db	2813	AGAAAGTTAAATACATGATACAGAGAGTTTGCAAGTACATACAGAAATCTTTTGAAGA	2872
QY	2823	AAGCAACAATCTTGGGGGTCTCATTTTCTATACCCAGGCAATGAGTATTTTTCCTATG	2882
Db	2873	AAGCAACAATCTTGGGGGTCTCATTTTCTATACCCAGGCAATGAGTATTTTTCCTATG	2932
QY	2883	CTGAGCTGTTTCCGGTTGGTGGCTACTTGGTGGCAATAGTTTCACTTGACAGATG	2942
Db	2933	CTGAGCTGTTTCCGGTTGGTGGCTACTTGGTGGCAATAGTTTCACTTGACAGATG	2992
QY	2943	TTCTTTTGGATTTCTCAGATATGTCTTTGGTGGCACTGGCAATGGGGAGGTCAGTTTAT	3002
Db	2993	TTCTTTTGGATTTCTCAGATATGTCTTTGGTGGCACTGGCAATGGGGAGGTCAGTTTAT	3052

QY	3003	TTGCTCTGACTATGCCAAAGCCAAAGTATACAGAGCCCGCATCATGATCATTTGA	3065
Db	3053	TTGCTCTGACTATGCCAAAGCCAAAGTATACAGAGCCCGCATCATGATCATTTGA	3112
QY	3063	AAAGCCCTCATATTGACAGCTACAGCCCTAGAGCCCTCAAGCCAAATACGTGGAAGAA	3122
Db	3113	AAAGCCCTCATATTGACAGCTACAGCCCTACGCGCTCAAGCCAAATACGTGGAAGAA	3172
QY	3123	ATGTGACATTTAATGAGGTGTGTCAACATCCGCTGACACACATCCCGCTGCTCC	3182
Db	3173	ATGTGACATTTAATGAGGTGTGTCAACATCCGCTGACACACATCCCGCTGCTCC	3232
QY	3183	AGGGGCTGACCCCTGAGGTGTAAGAAGGGCCAGACGCTGGCCCTCTGATGTAAGTGGC	3242
Db	3233	AGGGGCTGACCCCTGAGGTGTAAGAAGGGCCAGACGCTGGCCCTCTGATGTAAGTGGC	3292
QY	3243	GTCGGAAGACACAGTTGTTACAGCTCTGTGAGCGCTTCTATGACCCTTGAGCTGTTG	3302
Db	3293	GTCGGAAGACACAGTTGTTACAGCTCTGTGAGCGCTTCTATGACCCTTGAGCTGTTG	3352
QY	3303	TTCGTAATGTGATGCCAAAGAGATTAAGCACCTGTAATGTCAGTGGCTCCGAGCACACTGG	3362
Db	3353	TTCGTAATGTGATGCCAAAGAGATTAAGCACCTGTAATGTCAGTGGCTCCGAGCACACTGG	3412
QY	3363	GCATCGTGTCTGAGGAGCCCATCCGTTTGACTGACAGCATTTGCCGAGACATTCCTATG	3422
Db	3413	GCATCGTGTCTGAGGAGCCCATCCGTTTGACTGACAGCATTTGCCGAGACATTCCTATG	3472
QY	3423	GAGACACAGCCGGGTCGTATACATGTAAGAATTAATGACAGGACCAAGAGGCCCA	3482
Db	3473	GAGACACAGCCGGGTCGTATACATGTAAGAATTAATGACAGGACCAAGAGGCCCA	3532
QY	3483	TACACACACTCATCTGACAGCACTCCCTGAGAAATTAACAAACACAGTAAGAACAAGAA	3542
Db	3533	TACACACACTCATCTGACAGCACTCCCTGAGAAATTAACAAACACAGTAAGAACAAGAA	3592
QY	3543	CCGACCTCTCTGTGTGCCAGAAACAGCGCATTCGATGCTGCCTCTGTTAGACAGC	3602
Db	3593	CCGACCTCTCTGTGTGCCAGAAACAGCGCATTCGATGCTGCCTCTGTTAGACAGC	3652
QY	3603	CTCATATTTTCTTTGGATGTAAGCTACATCACCCTCGATACAGAAAGTGAAGGTTG	3662
Db	3653	CTCATATTTTCTTTGGATGTAAGCTACATCACCCTCGATACAGAAAGTGAAGGTTG	3712
QY	3663	TCCAAAGAGCCCTGTGCACAAAGCCAGAGAGCCGACCTGCATTTGTGATGCCACCGCT	3722
Db	3713	TCCAAAGAGCCCTGTGCACAAAGCCAGAGAGCCGACCTGCATTTGTGATGCCACCGCT	3772
QY	3723	TGTCACACATCCGAAATGCAAGATTTAATAGTGTGTTTCAATATGCAATGCAAGAGC	3782
Db	3773	TGTCACACATCCGAAATGCAAGATTTAATAGTGTGTTTCAATATGCAATGCAAGAGC	3832
QY	3783	ATGGCACACATCAACAGCTGTGTGCCAGAAAGGATTTATTTTCCATGCTCAAGTCTC	3842
Db	3833	ATGGCACACATCAACAGCTGTGTGCCAGAAAGGATTTATTTTCCATGCTCAAGTCTC	3892
QY	3843	AGGCTGGAGAAAGGCGCTGTCACTGTGTGCCATATGAGCTGTTAAATATTTTATAT	3902
Db	3893	AGGCTGGAGAAAGGCGCTGTCACTGTGTGCCATATGAGCTGTTAAATATTTTATAT	3952
QY	3903	TGTGTATTAACATGGCATTTAATCAAAAGTTAAAGGAGCACTTAATGAGAAACAT	3962
Db	3953	TGTGTATTAACATGGCATTTAATCAAAAGTTAAAGGAGCACTTAATGAGAAACAT	4012
QY	3963	GATAGAACTACCTGTTTAACATTTCTGTGCTGCACTGAAGTCAATTCACCAAGTTCAAG	4022
Db	4013	GATAGAACTACCTGTTTAACATTTCTGTGCTGCACTGAAGTCAATTCACCAAGTTCAAG	4072
QY	4023	TCTTGAATTTTATTAATTAAGAAACCAAAAGAAACATTAATGATGGAATTAATATG	4082
Db	4073	TCTTGAATTTTATTAATTAAGAAACCAAAAGAAACATTAATGATGGAATTAATATG	4132
QY	4083	GTGTAAATGCAATTAATAATTTATAGAGTAATTCAAAGTAGATTTTGTTAATTAATTTGTA	4142

Query Match 89.9% Score 3847.4 DB 12: Length 4045;	
Best Local Similarity 99.7% Pred. No. 0;	
Matches 3854; Conservative 0; Mismatches 11; Indels 0; Gaps	
QY	20 GATCCTTAAGAGAGCCGCTAAGGGAGTCGACAGAAAGAACTTCTGGAAATATGGCAAAAA 79
DB	122 GATCCTTAAGAGAGCCGCTAAGGGAGTCGACAGAAAGAACTTCTGGAAATATGGCAAAAA 181
QY	80 AGTAAAAAATAGAGAGAAAGAAAAGAAACAACGTGACGACGTTTGCAATGTTTGGC 139
DB	182 AGTAAAAAATAGAGAGAAAGAAAAGAAACAACGTGACGACGTTTGCAATGTTTGGC 241
QY	140 TATTCAAAATGGGTTATATAGTGTATATGTGTGGGAGCAATAGGTCCTATCATCAT 199
DB	242 TATTCAAAATGGGTTATATAGTGTATATGTGTGGGAGCAATAGGTCCTATCATCAT 301
QY	200 GGAGCTGACATCCCTCTCATATGCTGTGTTTGGAAACATGACGATAGCTTTGCCAAAT 259
DB	302 GGAGCTGACATCCCTCTCATATGCTGTGTTTGGAAACATGACGATAGCTTTGCCAAAT 361
QY	260 GCAGGAATTTTCAAGAAACAAACTTTTCCAGTATATATTAATGAAGATATACGACAAAT 319
DB	362 GCAGGAATTTTCAAGAAACAAACTTTTCCAGTATATATTAATGAAGATATACGACAAAT 421
QY	320 ACACAATATTTTCATCAACCATCTGAGAGAGAAATACCAAGTATGCTATATTAACAT 379
DB	422 ACACAATATTTTCATCAACCATCTGAGAGAGAAATACCAAGTATGCTATATTAACAT 481
QY	380 GGAATGCGTCTGCGCGTGTGCTGGGCGTCTAAATCCAGTTTCAATTCGTGCGCTGGCA 439
DB	482 GGAATGCGTCTGCGCGTGTGCTGGGCGTCTAAATCCAGTTTCAATTCGTGCGCTGGCA 541
QY	440 GCAGGAAGACAGATACCAAAATTAGAAAACAATTTTTCATGCTATATCAGACGAGAG 499
DB	542 GCAGGAAGACAGATACCAAAATTAGAAAACAATTTTTCATGCTATATCAGACGAGAG 601
QY	500 ATTGGCTGGTTTACGTGCATGACGTTGGGAGACTTAACACCGGCTACAGACAGATGTC 559
DB	602 ATTGGCTGGTTTACGTGCATGACGTTGGGAGACTTAACACCGGCTACAGACAGATGTC 661
QY	560 TCCAAAATCAATGAGAAATTTGGGCAAAAATTGGAAATGTTTTCACATATAGCAACA 619
DB	662 TCCAAAATCAATGAGAAATTTGGGCAAAAATTGGAAATGTTTTCACATATAGCAACA 721
QY	620 TTTTTCACGGGTTTATAGTGGGGTTTACAAGTGTGTGGAACCTAACCTTGATATTTTG 679
DB	722 TTTTTCACGGGTTTATAGTGGGGTTTACAAGTGTGTGTGGAACCTAACCTTGATATTTTG 781
QY	680 GCATCAGCCGCTTTCTTGGACTTTTCAGCCGCCATCTGGGGCAAAATATCTATCTCATTT 739
DB	782 GCATCAGCCGCTTTCTTGGACTTTTCAGCCGCCATCTGGGGCAAAATATCTATCTCATTT 841
QY	740 ACTGATTAAGAACTCTGTGGCCTATGCAAAAAGCTGAGACAGTATGCTGAAGAAGCTTAGCA 799
DB	842 ACTGATTAAGAACTCTGTGGCCTATGCAAAAAGCTGAGACAGTATGCTGAAGAAGCTTAGCA 901
QY	800 GCATCAGAACTGTATGTCCTTTTGGAGCAAAAAGAAAGAACTTGAAGAGTACACACAA 859
DB	902 GCATCAGAACTGTATGTCCTTTTGGAGCAAAAAGAAAGAACTTGAAGAGTACACACAA 961
QY	860 AATTTGAGAGAGCTTAAGGAAATTTGGATTAAGAAAGAAAGTACAGGCCACATTTCTATT 919
DB	962 AATTTGAGAGAGCTTAAGGAAATTTGGATTAAGAAAGAAAGTACAGGCCACATTTCTATT 1021
QY	920 GGGGCCGCTTTCTTATATATCTATGCAATCATATGCTGTGCTTCTGTATATGGACCTCC 979
DB	1022 GGGGCCGCTTTCTTATATATCTATGCAATCATATGCTGTGCTTCTGTATATGGACCTCC 1081
QY	980 TTGGTCTCTCCAGTGAATATTTATTTGACACAGTACTACCTGCTTCTTTCTGTATTA 1035
DB	1082 TTGGTCTCTCCAGTGAATATTTATTTGACACAGTACTACCTGCTTCTTTCTGTATTA 1141

	gene	/tissue_type="liver"	/note="Ivermectin sensitive collies"
	CDS	7..282 /gene="p-gp" /gene="p-gp" /name="abortive product" /evidence=experimental /product="p-glycoprotein" /protein_id="CAD12806_1" /db_xref="GI:17385399" /translation="MDPEGRGRRGSAEKNEFKMKQSKKKNEKKKPYTSFPMFRYS WIDRYIMVGTMAALIHGAALPLMLVLFGMTALDOMEFOETRLRQ"	
	BASE COUNT	1144 a 815 c 964 g 1011 t	
	ORIGIN		
	Query Match	89.4%; Score 3824; DB 4; Length 3934;	
	Best Local Similarity	99.4%; Pred. No. 0;	
	Matches 3850; Conservative	0; Mismatches 20; Indels 4; Gaps 1	
OY	11 GTCCGGATGCATCTCTGAAGAGGCCGTAAAGGGAGTGCAGACAGACTTTGGAAAATG 70		
Dd	1 GTGACAGTNGATCTCTGAAGAGGCCGTAAGGGAGTGCGAGAAACAACCTTGTGAAGAATG 60		
OY	71 GGCAAAAAAAGTAAAAAAATGAGAGAAAGAAAACCAACTGTGACAGCTTTGCA 130		
Dd	61 GGCAAAAAAAGTAAAAAAATGAGAGAAAGAAAACCAACTGTGACAGCTTTGCA 120		
OY	131 AGTTTCGCTATTCAAATTTGGCTTGATAGTGTATATGTTGGTGGGCAATGGCTGCC 190		
Dd	121 ATGTTTCGCTATTCAAATTTGGCTTGATAGTGTATATGTTGGTGGGCAATGGCTGCC 180		
OY	191 ATCATCCATGAGAGCTGCACCTCCTCTCATATGCTGGTTTTGGAAACATGACAGATAGC 250		
Dd	181 ATCATCCATGAGAGCTGCACCTCCTCTCATATGCTGGTTTTGGAAACATGAC---AGC 236		
OY	251 TTTTGCAAATGCAGAAATTTCAAGAAACAAAATTTTCCAGTTATTAATTAAGAAATAT 310		
Dd	237 TTTTGCAAATGCAGAAATTTCAAGAAACAAAATTTTCCAGTTATTAATTAAGAAATAT 296		
OY	311 ACGAACATATACAAACAATTTTCATCAACCACCTGGAGAGAAATGACCAGATAGCCTAT 370		
Dd	297 ACGAACATATACAAACAATTTTCATCAACCACCTGGAGAGAAATGACCAGATAGCCTAT 356		
OY	371 TATTACAGTAGGAGTCGGTGGTGGCTGGCTGCTGCTCTTACATCCAGTTTCATTTCTGG 430		
Dd	357 TATTACAGTAGGAGTCGGTGGTGGCTGGCTGCTGCTCTTACATCCAGTTTCATTTCTGG 416		
OY	431 TGCCCTGGACAGCAGAGAACAGATACTCCAATAATTTGAAAACATTTTTTCTCAGTATCATG 490		
Dd	417 TGCCCTGGACAGCAGAGAACAGATACTCCAATAATTTGAAAACATTTTTTCTCAGTATCATG 476		
OY	491 CGACAGAGAAATGGCTGGTTTACAGTGACGACGTTGGGAGCTTAACACCCGGCTCAC 550		
Dd	477 CGACAGAGAAATGGCTGGTTTACAGTGACGACGTTGGGAGCTTAACACCCGGCTCAC 536		
OY	551 GACGATGTCTCCAAATCAATGAAGAAATTTGGCGCAAAAATTTGGAATGATCTTTCACTCA 610		
Dd	537 GACGATGTCTCCAAATCAATGAAGAAATTTGGCGCAAAAATTTGGAATGATCTTTCACTCA 596		
OY	611 ATTACACATTTTTTCAACCGGTTTATATAGTGGGGTTTACACGTTGGAGACTTAACCTT 670		
Dd	597 ATTACACATTTTTTCAACCGGTTTATATAGTGGGGTTTACACGTTGGAGACTTAACCTT 656		
OY	671 GTGATTTTGGCATCAGCCCTGTTCTTGGACTTTCAAGCCGCATGCTGGGCAAAATACTA 730		
Dd	657 GTGATTTTGGCATCAGCCCTGTTCTTGGACTTTCAAGCCGCATGCTGGGCAAAATACTA 716		
OY	731 TCTTCATTACTAGTAAGAAGCTTTGGCCTATGCAAAAAGCTGGAAGCAGTAGCTGAAGAA 790		
Dd	717 TCTTCATTACTAGTAAGAAGCTTTGGCCTATGCAAAAAGCTGGAAGCAGTAGCTGAAGAA 776		


```
QY 791 GTCTTAGAGCAATACAGACACTGATTCGCTTTGGAGAGCAAAAGAAAGAACTGAAAG 850
Db 777 GTCTTAGAGCAATACAGAACTGATTCGCTTTGGAGAGCAAAAGAAAGAACTGAAAG 836
QY 891 TACAGCAAAATTTAGAAAGAGCTAAAGAAATGGGATTAAGAAAGCTATCACGGCCAC 910
Db 837 TACAGCAAAATTTAGAAAGAGCTAAAGAAATGGGATTAAGAAAGCTATCACGGCCAC 896
QY 911 ATTTCTATTTGGTCCGCTTTCTATGATCTATGATCATCATGCTGCTGCTTCTGAT 970
Db 897 ATTTCTATTTGGTCCGCTTTCTATGATCTATGATCATCATGCTGCTGCTTCTGAT 956
QY 971 GGGACCTCCTTGCTCTCCAGAGTAATTCCTATTTGGACAAGTACTGCTCTTTT 1030
Db 957 GGGACCTCCTTGCTCTCCAGAGTAATTCCTATTTGGACAAGTACTGCTCTTTT 1016
QY 1031 TCTGTATTTATTTGGGCTTTTGTATTTGACAGGATCCCAAGCATTTGAACATTTGCA 1090
Db 1017 TCTGTATTTATTTGGGCTTTTGTATTTGATGACAGGATCCCAAGCATTTGAACATTTGCA 1076
QY 1091 AACGCAAGAGAGACACTTATGAATCTTCAAGATATTTGACAAATTAACCAAGCATTGAC 1150
Db 1077 AACGCAAGAGAGAGACACTTATGAATCTTCAAGATATTTGACAAATTAACCAAGCATTGAC 1136
QY 1151 AGCTATTTGAAGAGTGGACATTAACCAAGATATTTAAAGGAAATTTGCAATTCAAAAAT 1210
Db 1137 AGCTATTTGAAGAGTGGACATTAACCAAGATATTTAAAGGAAATTTGCAATTCAAAAAT 1196
QY 1211 GTTCACTTCACTTACCTTCTCGAAGAAAGATTAAGATCTTAAAGGCTTCACCTGAC 1270
Db 1197 GTTCACTTCACTTACCTTCTCGAAGAAAGATTAAGATCTTAAAGGCTTCACCTGAC 1256
QY 1271 GTTCAAGAGTGGGACAGAGTGGCGCTGTTGGAAACAGTGGCTCGGGAGAGACAGACAC 1330
Db 1257 GTTCAAGAGTGGGACAGAGTGGCGCTGTTGGAAACAGTGGCTCGGGAGAGACAGACAC 1316
QY 1331 GTGACAGTGTAGAGAGGCTCTATGACCCACAGATGGCATGGCTCTATTTGATGACAG 1390
Db 1317 GTGACAGTGTAGAGAGGCTCTATGACCCACAGATGGCATGGCTCTATTTGATGACAG 1376
QY 1391 GACATTAAGACCAATTAATGTAAGGCATCTTCGGGAAATTACTGCTGTGTGATGACAGAG 1450
Db 1377 GACATTAAGACCAATTAATGTAAGGCATCTTCGGGAAATTACTGCTGTGTGATGACAGAG 1436
QY 1451 CCTGTGTTGTTTGGCACACAGATAGCTGGAAGAAATTCGCTATGCGCGGAAATGTCACC 1510
Db 1437 CCTGTGTTGTTTGGCACACAGATAGCTGGAAGAAATTCGCTATGCGCGGAAATGTCACC 1496
QY 1511 ATGATGATGATGGAAGAGCTGTTAAGAGAACCAATGCTATGATTTTATCATGAAACTA 1570
Db 1497 ATGATGATGATGGAAGAGCTGTTAAGAGAACCAATGCTATGATTTTATCATGAAACTA 1556
QY 1571 CCTAATTAATTTTGACACTCTGCTGTTGGAGAGAGAGGCGCCAGCTGAGTGGTGAAGAA 1630
Db 1557 CCTAATTAATTTTGACACTCTGCTGTTGGAGAGAGAGGCGCCAGCTGAGTGGTGAAGAA 1616
QY 1631 CAGAGAAATCCCAATGCTGCGGGCCCTGCTGTTGGAACCCCAAGATTTCTGCTGATGAG 1690
Db 1617 CAGAGAAATCCCAATGCTGCGGGCCCTGCTGTTGGAACCCCAAGATTTCTGCTGATGAG 1676
QY 1691 GCMAAGTACACTCTGACAGCTGAAAGTGAAGAGTGGTTCAGTGGCGCTGGATTAAGGCC 1750
Db 1677 GCMAAGTACACTCTGACAGCTGAAAGTGAAGAGTGGTTCAGTGGCGCTGGATTAAGGCC 1736
QY 1751 AGAAAAGCGGAGACTACCAATTTGATAGTCAATCGTTTGTCTACAGTTCGTAATGCCAT 1810
Db 1737 AGAAAAGCGGAGACTACCAATTTGATAGTCAATCGTTTGTCTACAGTTCGTAATGCCAT 1796
QY 1811 GTGATTCGTGTTTTGATGATGAGTCAATTTGAGAGAAAGAAATCATGAGACTCATG 1870
Db 1797 GTGATTCGTGTTTTGATGATGAGTCAATTTGAGAGAAAGAAATCATGAGACTCATG 1856
QY 1871 AAAGAGAAAGGAGATTACTTCAAACTGTGCACAAATGACAGACAAGAGAAATGAATTGAG 1930
Db 1857 AAAGAGAAAGGAGATTACTTCAAACTGTGCACAAATGACAGACAAGAGAAATGAATTGAG 1916
QY 1931 TTGAAAAATGCCAGTGTGATTCGAAAGTGAAGTATGCTTGGAAATGTCCTCAAAA 1990
Db 1917 TTGAAAAATGCCAGTGTGATTCGAAAGTGAAGTATGCTTGGAAATGTCCTCAAAA 1976
QY 1991 GATTCAAGGCTCCAGTTTAATTAAGAAATCAACCTCGAGAGATATCATGCAACCAAA 2050
Db 1977 GATTCAAGGCTCCAGTTTAATTAAGAAATCAACCTCGAGAGATATCATGCAACCAAA 2036
QY 2051 GGGCAAGACAGAAAGCTTGTGTCAAAAGAGACTGAATGAGAAATGACTCCAGTTTCC 2110
Db 2037 GGGCAAGACAGAAAGCTTGTGTCAAAAGAGACTGAATGAGAAATGACTCCAGTTTCC 2096
QY 2111 TTCTGAGAGATTCCTGAAGCTGAAGTCACTGAATGAGCCCTTATTTGTGTTGTAATTT 2170
Db 2097 TTCTGAGAGATTCCTGAAGCTGAAGTCACTGAATGAGCCCTTATTTGTGTTGTAATTT 2156
QY 2171 TGTGCTATTTAATTAACGGAGGCTGCAACAGATTTCAATTAATTTTCAAGATTAATA 2230
Db 2157 TGTGCTATTTAATTAACGGAGGCTGCAACAGATTTTCAATTAATTTTCAAGATTAATA 2216
QY 2231 GGGATCCTTAACCCGAGATGAGGATCTGAAACAAAGACAGAAATAGTAACATGTTTCT 2290
Db 2217 GGGATCCTTAACCCGAGATGAGGATCTGAAACAAAGACAGAAATAGTAACATGTTTCT 2276
QY 2291 GTATTGTTTCAAGTCTTGAATATTTCTTTATTTATTTTCCGCGAGGCTTCACA 2350
Db 2277 GTATTGTTTCAAGTCTTGAATATTTCTTTATTTATTTTCCGCGAGGCTTCACA 2336
QY 2351 TTTGGCAAGCTGGGAGATCTCTACTAAGCGGCTTGATACATGTTTTCAGATCCATG 2410
Db 2337 TTTGGCAAGCTGGGAGATCTCTACTAAGCGGCTTGATACATGTTTTCAGATCCATG 2396
QY 2411 CTGACACAGATGTCAGCTGTTGATGACCCCTTAAGAAACCACTGAGAGATTTGACACC 2470
Db 2397 CTGACACAGATGTCAGCTGTTGATGACCCCTTAAGAAACCACTGAGAGATTTGACACC 2456
QY 2471 AGGCTTGCCAAATGATGCGGCTCAAGTTAAAGGGCTATAGTTCAGAGCTGCTGCATTT 2530
Db 2457 AGGCTTGCCAAATGATGCGGCTCAAGTTAAAGGGCTATAGTTCAGAGCTGCTGCATTT 2516
QY 2531 ACCGAAATATAGCAATCTTGGACAGGCAATATATATCTTATTCATGTTGGCAA 2590
Db 2517 ACCGAAATATAGCAATCTTGGACAGGCAATATATATCTTATTCATGTTGGCAA 2576
QY 2591 TTAACACTTTTACTCTTAGCAATTTGTAACCATCATTTGCAATGACAGAGTGTGAAATG 2650
Db 2577 TTAACACTTTTACTCTTAGCAATTTGTAACCATCATTTGCAATGACAGAGTGTGAAATG 2636
QY 2651 AAAATGTTGCTGACAGAGCTGAAAGATTAAGAAAGACTTGAAGAGAGCTGGAAAGTT 2710
Db 2637 AAAATGTTGCTGACAGAGCTGAAAGATTAAGAAAGACTTGAAGAGAGCTGGAAAGTT 2696
QY 2711 GCTACAGAACCCATCGAAAACTTCGAACCTGTTTCTTATGACGAGGAGAGAAAGTT 2770
Db 2697 GCTACAGAACCCATCGAAAACTTCGAACCTGTTTCTTATGACGAGGAGAGAAAGTT 2756
QY 2771 GAATTAATGATGACAGAGTGTGCAAGTACATACAGAAACTTTTGAAGAAAGCACAC 2830
Db 2757 GAATTAATGATGACAGAGTGTGCAAGTACATACAGAAACTTTTGAAGAAAGCACAC 2816
QY 2831 ATCTTCGGGGTCTCATTTCTATCACCCGAGCAATGATTTTCTCTATGCGGCTGT 2890
Db 2817 ATCTTCGGGGTCTCATTTCTATCACCCGAGCAATGATTTTCTCTATGCGGCTGT 2876
QY 2891 TTCGGTTTGTGCTCTACTGTTGACAAATGATTCATGAACCTTTCAGAGATGTTCTTTG 2950
Db 2877 TTCGGTTTGTGCTCTACTGTTGACAAATGATTCATGAACCTTTCAGAGATGTTCTTTG 2936
QY 2951 GTATTCTCAGCTATTTGCTTTGTTGGCATGCGAGTGGGAGGTCAGTTTCTGCTCT 3010
Db 2937 GTATTCTCAGCTATTTGCTTTGTTGGCATGCGAGTGGGAGGTCAGTTTCTGCTCT 3010
```

Db 2937 GATATTCAGCTATGCTTCTTGGTGCATGCGAGTGGGGGAGGTGAGTTCATTGCTCCT 2996
Oy 3011 GACTATGCCAAAGCCAAAGATATGACAGACCCAGCTCATGATCATTTAAAAAAGCCCT 3070
Db 2997 GACTATGCCAAAGCCAAAGATATGACAGACCCAGCTCATGATCATTTAAAAAAGCCCT 3056
Oy 3071 CTGATTCAGCTATGAGCCCTCAGGCTCAGCAAAATGCTTGGAGAAATGTGACA 3130
Db 3057 CTGATTCAGCTATGAGCCCTCAGGCTCAGCAAAATGCTTGGAGAAATGTGACA 3116
Oy 3131 TTTATGAGCTGCTGTTCACTATCCACTGACAGACATCCCGTCTCCAGGGCTG 3190
Db 3117 TTTATGAGCTGCTGTTCACTATCCACTGACAGACATCCCGTCTCCAGGGCTG 3176
Oy 3191 AGCCTCGAGTGAAGAGAGGCCCAAGCTGGCCCTCTAGTGTGAGTGGCTGGGAAG 3250
Db 3177 AGCCTCGAGTGAAGAGAGGCCCAAGCTGGCCCTCTAGTGTGAGTGGCTGGGAAG 3236
Oy 3251 AGCAGATGTTCACTCTCTAGAGCGCTTCTATGACCCCTTGGCTGGTCACTATAT 3310
Db 3237 AGCAGATGTTCTAGCTCTCTAGAGCGCTTCTATGACCCCTTGGCTGGTCACTATAT 3296
Oy 3311 GATGGCAAGAGATTAAGACACTGATATGTCAGTGGCTCCGAGACACCTGGGATGCTG 3370
Db 3297 GATGGCAAGAGATTAAGACACTGATATGTCAGTGGCTCCGAGACACCTGGGATGCTG 3356
Oy 3371 TCTCAGAGACCCATCTGTTTACATGACAGATTCGCGAATTCCTATGAGACAC 3430
Db 3357 TCTCAGAGACCCATCTGTTTACATGACAGATTCGCGAATTCCTATGAGACAC 3416
Oy 3431 AGCCGGCTGATTCACATGAGATTAATGACAGAGCAGAGAGAGGCCCAATCATCACAC 3490
Db 3417 AGCCGGCTGATTCACATGAGATTAATGACAGAGCAGAGAGAGGCCCAATCATCACAC 3476
Oy 3491 TTGATGAGACCTCCCTGAGAAATACAAACACAGATGAGAGAGAGAGAGAGAGAGAG 3550
Db 3477 TTGATGAGACCTCCCTGAGAAATACAAACACAGATGAGAGAGAGAGAGAGAGAGAG 3536
Oy 3551 TCTGCTGGCCAGAAAGAGCGCATTCGATGCTGCGCTCTTGTGTAGACAGCTCATATT 3610
Db 3537 TCTGCTGGCCAGAAAGAGCGCATTCGATGCTGCGCTCTTGTGTAGACAGCTCATATT 3596
Oy 3611 TTGCTTTGGATTAAGACTCATAGCTGATGATGATGATGATGATGATGATGATGATGAT 3670
Db 3597 TTGCTTTGGATTAAGACTCATAGCTGATGATGATGATGATGATGATGATGATGATGAT 3656
Oy 3671 GCCCTGGACAAAGCCAGAGAGAGCGCATTCGATGCTGCGCTCTTGTGTAGACAGCTCATATT 3730
Db 3657 GCCCTGGACAAAGCCAGAGAGAGCGCATTCGATGCTGCGCTCTTGTGTAGACAGCTCATATT 3716
Oy 3731 ATCCAGATGCAATTAATAGTGTGTTTCAGATGCAAGAGTCAAGAGATGGACACA 3790
Db 3717 ATCCAGATGCAATTAATAGTGTGTTTCAGATGCAAGAGTCAAGAGATGGACACA 3776
Oy 3791 CATCAACAGCTGCTGCCAGAGAGAGCATTTTTCATGCTGATGCTGATGCTGATGCTGAT 3850
Db 3777 CATCAACAGCTGCTGCCAGAGAGAGCATTTTTCATGCTGATGCTGATGCTGATGCTGAT 3836
Oy 3851 GCAAAGCGTATGATGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 3884
Db 3837 GCAAAGCGTATGATGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 3870

RESULT 9
LOCUS 108557 108557 4669 bp DNA linear PAT 02-DEC-1994
DEFINITION Sequence 3 from Patent WO 8705943.
ACCESSION 108557
VERSION 108557.1 GI:588735
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
Unclassified.

REFERENCE 1 (bases 1 to 4669)
AUTHORS Roninson, I. B., Pastan, I. H. and Gottesman, M. M.
TITLE COMPOSITIONS AND METHODS FOR CLONES CONTAINING DNA SEQUENCES
JOURNAL ASSOCIATED WITH MULTIDRUG RESISTANCE IN HUMAN CELLS
PATENT Patent: WO 8705943-A 3 08-OCT-1987;
FEATURES Location/Qualifiers
source 1..4669
BASE COUNT 1394 a 892 c 1129 g 1254 t
ORIGIN
Query Match 78.1%; Score 3442.8; DB 6; Length 4669;
Best Local Similarity 87.6%; Pred. No. 0;
Matches 3759; Conservative 0; Mismatches 487; Indels 44; Gaps 8;
Oy 1 GGAGCGGAGGTGGGATGATGATCTGTAAGAGAGCCCTTAAGGGAGAGTGCAGAGAACTT 60
Db 409 GGAGCGGAGGTGGGATGATGATCTGTAAGAGAGCCCTTAAGGGAGAGTGCAGAGAAAGAA 468
Oy 61 CTGGAATAATGGGCAAAAAAGTAAAAAATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 120
Db 469 CTTTCTTAATGAAACATTAATAAGTGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 528
Oy 121 CACGTTTCATGTTTGGCTTTCATTAATGCTGATAGGTGATATGTTGGGGAGAC 180
Db 529 TGTATTTCAATGTTTCCCTATCAAAATGCTGACAGATGATATGTTGGTGGAGAC 588
Oy 181 AATGGCTGCATCATCATGAGAGCTGCTCCTCATGATGATGCTGTTTGGAAACAT 240
Db 589 TTGGCTGCATCATCATGAGAGCTGCTCCTCATGATGATGCTGTTTGGAAACAT 648
Oy 241 GACAGATACCTTTCGAAATGACAGAAATTCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 300
Db 649 GACAGATACCTTTCGAAATGACAGAAATTCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 702
Oy 301 TGAAGTATTAAGAACAAATACACATTCATGACCATCTGAGAGAGAGAGAGAGAGAGAG 360
Db 703 TAATGAAGTATTAAGAACAAATACACATTCATGACCATCTGAGAGAGAGAGAGAGAGAG 762
Oy 361 GATATGCTATTAATACAGTGGAGTGGTGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 420
Db 763 GATATGCTATTAATACAGTGGAGTGGTGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 822
Oy 421 TTGATTCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 480
Db 823 TTGATTCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 882
Oy 481 TGTATCATGCGACAGAGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 540
Db 883 TGTATCATGCGACAGAGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 942
Oy 541 CCGGCTCAGAGAGATGCTCCAAATCAATGAAGAAATGGCGCAAAATTTGCAATGTT 600
Db 943 CCGGCTCAGAGAGATGCTCCAAATCAATGAAGAAATTTGCGCAAAATTTGCAATGTT 1002
Oy 601 CTTTCACTCAATGAACAAATTTTCAACCGTCTTATAGTGGGTTTACACGTGTTGCA 660
Db 1003 CTTTCACTCAATGAACAAATTTTCAACCGTCTTATAGTGGGTTTACACGTGTTGCA 1062
Oy 661 GCTAACCTTGATTTTGGCCATCAGCCCTGCTTGGAGACTTACGCCGCTGAGGCTG 720
Db 1063 GCTAACCTTGATTTTGGCCATCAGCCCTGCTTGGAGACTTACGCCGCTGAGGCTG 1122
Oy 721 AAAGTACTATCTTCAATTAAGTAAGAACTTGGCTATGCAAAAGCTGAGCACT 780
Db 1123 AAAGTACTATCTTCAATTAAGTAAGAACTTGGCTATGCAAAAGCTGAGCACT 1182
Oy 781 AGCTGAAGAGTCTTACAGCAATCAGAACTGATTCCTTTGGAGGACAAAGAGAA 840
Db 1183 AGCTGAAGAGTCTTGGCAGCAATTAAGAACTGATTCCTTTGGAGGACAAAGAGAA 1242
Oy 841 ACTGAAAGTACAACAAAATTTAGAGAAAGCTAAAGAAATGGATTAAGAAAGCTAT 900
Db 1242 ACTGAAAGTACAACAAAATTTAGAGAAAGCTAAAGAAATGGATTAAGAAAGCTAT 900

Db 1243 ACTGGAAGGTACACAAATAATTTAGAGAAGCTAAAGAAATGGGATAAAGAAAGCTAT 1302
QY 901 CACGGCCAAACATTTCTATGTGGCCGCTTCTTATGTATGTATGCATATATGCTTGCG 960
Db 1303 TACAGCCAAATTTCTATGTAGGTGCTGCTTCCGCTGATCTATGATCTTATGCTTGCG 1362
QY 961 TTTTGTGTATGGAGACTCTCTTGGTCTCTCCAGTGAATATTTATTTGGCAAGTACTAC 1020
Db 1363 CTTCGTGTATGGACCACTTGTCTCTCAGGGGAATTTCTATTTGGACAAAGTACTAC 1422
QY 1021 TGTCTTTCTTCTGTATTTATTTGGGCTTTTATGTATTTGGACAGGATCCCAAGCATTA 1080
Db 1423 TGTATCTTTTCTGTATTTATTTGGGCTTTTATGTATTTGGACAGGATCTCCAAAGCATTA 1482
QY 1081 AGCATTTGCAAAACGCAAGAGAGAGACCTTATGAAATCTTCAAGATTAATTTGACATTAAC 1140
Db 1483 AGCATTTGCAAAATGCAAGAGAGAGACCTTATGAAATCTTCAAGATTAATTTGAAATTAAC 1542
QY 1141 AAGCATTTGACAGCTATTTCCAAAGGTGACATTAACCAAGATTAATTTAAGGAAATTTGA 1200
Db 1543 AAGTATTTGACAGCTATTTCCAAAGGTGACATTAACCAAGATTAATTTAAGGAAATTTGA 1602
QY 1201 AATTCAAAATGCTTCACTTACCTTCTTCCGAAAGAAAGTAAAGATCTTAAAGGCTCT 1260
Db 1603 AATTCAGAAATGCTTCACTTACCTTCTTCCGAAAGAAAGTAAAGATCTTAAAGGCTCT 1662
QY 1261 CAACCTGAAGGTTTCAGAGTGGGACAGACAGTGGCGCTGGTGGGAACAGTGGCTGGGAA 1320
Db 1663 GAACCTGAAGGTTTCAGAGTGGGACAGACAGTGGCGCTGGTGGGAACAGTGGCTGGGAA 1722
QY 1321 GAGCAGCAGCCGTGAGCTGATGAGAGGCTATGACCCCAACAGATGGCATTTGTAT 1380
Db 1723 GAGCAGCAACAGTCCAGCTGATGAGAGGCTATGACCCCAACAGAGGAGTGGTCACTGT 1782
QY 1381 TGATGCAAGAGCATTTAGGACCATTAATGTAAAGCATCTTGGGAAATTTAGTGGTGT 1440
Db 1783 TGATGCAAGAGCATTTAGGACCATTAATGTAAAGCATCTTGGGAAATTTAGTGGTGT 1842
QY 1441 GAGTCAGAGGCTGTGTTTGGCCACAGATAGCTGGAAGCAATTCGCTAGTGGCGCA 1500
Db 1843 GAGTCAGAGGCTGTGTTTGGCCACAGATAGCTGGAAGCAATTCGCTAGTGGCGCA 1902
QY 1501 AAATGTCCATGATGATGATGGAAGCTGTAAAGGCAAGCAATTCATGATTTAT 1560
Db 1903 AAATGTCCATGATGATGATGGAAGCTGTCAAGGAAGCAATTCATGATTTAT 1962
QY 1561 CATGAAATCTACTAATTAATTTGACACTCTGTTGAGAGAGAGGGGCCACAGTATG 1620
Db 1963 CATGAAATCTCTCTATTAATTTGACACCTGTTGGAGAGAGAGGGGCCACAGTATG 2022
QY 1621 TGGACAGAAACAGAGATGCCATTTGCTGGGGCCGTTGGCAACCCCAATCTCTCT 1680
Db 2023 TGGACAGAAACAGAGATGCCATTTGCTGGGGCCGTTGGCAACCCCAATCTCTCT 2082
QY 1681 GCTGAGAGAGGCAAGCTGACACTGTGACACTGAAGTGAAGCATGTTGACGTTGCGCT 1740
Db 2083 GCTGAGAGAGGCAAGCTGACACTGTGACACTGAAGTGAAGCATGTTGACGTTGCGCT 2142
QY 1741 GGATTAAGGCCAGAAAAGGCCGACTACATTTGTATAGCTATGCTTTGTCTACAGTTGC 1800
Db 2143 GGATTAAGGCCAGAAAAGGCCGACTACATTTGTATAGCTATGCTTTGTCTACAGTTGC 2202
QY 1801 TAATGCCAGTGCATTTGCTGTTTGTATGATGAGATCTTGTGAGAAAGAAATTCATGA 1860
Db 2203 TAATGCCAGTGCATTTGCTGTTTGTATGATGAGATCTTGTGAGAAAGAAATTCATGA 2262
QY 1861 TGAACCTCAGAAAGAGAGGAGATTTACTTCAAACTGTCTCAATGCAGACAGAGAAAA 1920
Db 2263 TGAACCTCAGAAAGAGAGGAGATTTACTTCAAACTGTCTCAATGCAGACAGAGAAAA 2322
QY 1921 TGAATTTAGATTAGAAAATGCCACTGTGTAATCCAAAAGTGAAGATGATGCTTGGAAAT 1980
Db 2323 TGAAGTTGAATTTAGAAAATGCAGCTGATGATCCAAAAGTGAATTTGATGCTTGGAAAT 2382

QY 1981 GTCTCCAAAAGATTTCAGAGGTCCAGTTTAAATTAAGAAGATCAACTCGAGAGATATCA 2040
Db 2383 GTCTCCAAAAGATTTCAGAGGTCCAGTTTAAATTAAGAAGATCAACTCGAGAGATATCA 2442
QY 2041 TGCACCAAGAGGCCCAAGACAGAAAGCTTGTACAAAAGAGGACTGTAATGAAATGTACC 2100
Db 2443 TGCATCCAAAGGCCCAAGACAGAAAGCTTGTACAAAAGAGGCTGTAATGAAATGTACC 2502
QY 2101 TCCAGTTTCTTCTGAGAGATTTCTGAAAGTCTGAAGTCAACTGAATGGCTTATTTGTGT 2160
Db 2503 TCCAGTTTCTTCTGAGAGATTTCTGAAAGTCTGAAGTCAACTGAATGGCTTATTTGTGT 2562
QY 2161 TGTATATTTTGTATTTATTAAGAGGCTGCAACACAGATTTTAAATATTTTC 2220
Db 2563 TGTATATTTTGTATTTATTAAGAGGCTGCAACACAGATTTTAAATATTTTC 2622
QY 2221 AAGGATTAAGGATCTTTACCGAGATGAGATCTGAAACAAAGCAGATATGAA 2280
Db 2623 AAGGATTAAGGATCTTTACCGAGATGAGATCTGAAACAAAGCAGATATGAA 2682
QY 2281 CATGTTTCTGATTTGTTTCTAGTCCCTGGAATATTTCTTTATACATTTTCTTCA 2340
Db 2683 CATGTTTCTGATTTGTTTCTAGTCCCTGGAATATTTCTTTATACATTTTCTTCA 2742
QY 2341 GGGCTTCAATTTGCAAAAGCTGGGAGATCTCTCAAGCGGCTTGATATACATGTTT 2400
Db 2743 GGGCTTCAATTTGCAAAAGCTGGGAGATCTCTCAAGCGGCTTGATATACATGTTT 2802
QY 2401 CAGATCCATGCTGAGACAGAGATGTCAGTGGTTGATGACCTTAAACACACTGGAG 2460
Db 2803 CAGATCCATGCTGAGACAGAGATGTCAGTGGTTGATGACCTTAAACACACTGGAG 2862
QY 2461 ATTGCAACCAAGGCTTGCAATGATGCGGCTCAAGTTAAAGGGCTATGAGTCCAGGCT 2520
Db 2863 ATTGCAACCAAGGCTTGCAATGATGCGGCTCAAGTTAAAGGGCTATGAGTCCAGGCT 2922
QY 2521 TGCCTGATTTCCCAAGAAATATGCAAAATCTTGGACAGCATTATTAATCTTAACTA 2580
Db 2923 TGCCTGATTTCCCAAGAAATATGCAAAATCTTGGACAGCATTATTAATCTTAACTA 2982
QY 2581 TGTGTGCAATTAACATTTTACTCTTAGCAATTTGACCATCATGATGCAATGACAGAGT 2640
Db 2983 TGTGTGCAATTAACATTTTACTCTTAGCAATTTGACCATCATGATGCAATGACAGAGT 3042
QY 2641 TGTGGAATGAATAATGTTGCTGACAGACACTGAAGAATGAAGAAGCTGGAAGAGC 2700
Db 3043 TGTGGAATGAATAATGTTGCTGACAGACACTGAAGAATGAAGAAGCTGGAAGAGC 3102
QY 2701 TGGGAAGATTTGCTACAGAAAGCCATGCAAAAATCTGTTGTTTCTTGTAGTCCGGA 2760
Db 3103 TGGGAAGATTTGCTACAGAAAGCCATGCAAAAATCTGTTGTTTCTTGTAGTCCGGA 3162
QY 2761 GCGAAGTTTGAATATCATGTATGACAGAGTTTGCAGTACCATCAGAAACTCTTTGAG 2820
Db 3163 GCGAAGTTTGAATATCATGTATGCTCAGAGTTTGCAGTACCATCAGAAACTCTTTGAG 3222
QY 2821 GAAAGCAGCATCTTCGGGCTCAATTTTCTATGCAACCCAGCATGATGATTTTCTTA 2880
Db 3223 GAAAGCAGCATCTTTGGATTTACATTTTCTTCCACCCAGCAATGATGATTTTCTTA 3282
QY 2881 TGTGTGCTTTTCCGTTTGTGTGCTACTTGTGTGCAATAGTTGATGAACTTTCAGA 2940
Db 3283 TGTGTGCTTTTCCGTTTGTGTGCTACTTGTGTGCAATAGTTGATGAACTTTCAGA 3342
QY 2941 TGTGTTTGTGATTTCTGAGCATTTGTGTTGTGTCATGAGCTTGGAG 3000
Db 3343 TGTGTTTGTGATTTTGTGCTGTTGTTGTGTCATGAGCTTGGAG 3402
QY 3001 ATTGCTCTGATATGCAAGGCAAGATTTACAGAGCCACGCTCATGATGATATGA 3060
Db 3403 ATTGCTCTGATATGCAAGGCAAGATTTACAGAGCCACATCTCATGATGATATGA 3462

OY	3061	AAAAAGCCCTCTGATTGACAGTCAAGCCCTCAAGCCCTCAAGCCAAATACGTGTGAAG	3120
Db	3463	AAAAACCCCTTTCGTTGATACAGCTACAGCACAGGAAGGCTTAATGCGAACAATTGGAAAG	3522
OY	3121	AAATGACATTTAATGAGGTCGCTCAACTATGCCACTCGACACAGACATCCCGCTGCT	3180
Db	3523	AAATGTCATTGTGGTAGAGTTGATTCACTATGCCACCAGCCGAGCATCCCAATGCT	3582
OY	3181	CCAGGGGCTGAAGCCCTCAGAGTGAAGAAGGGCCAGACGCTGGCCCTGGTAGGCTGAGTGG	3240
Db	3583	TCAGGAGTGTAGCCTGGAGGTGAAGAAGGGCCACACCTGGCTGTGGTGGGACAGACTGG	3642
OY	3241	CTGTGGGAAGGACAGAGTTGTTGAGCTCTAGAGACCGCTTTATGACCCCTTGGCTGGTTC	3300
Db	3643	CTGTGGGAAGGACAGAGTGGCTCCAGCTCTCGAGACCGGCTTCAGACCCCTTGGCAGGGAA	3702
OY	3301	AGTGTAAATTGATGGCAAAAGATTAAGCAACCTTAATGTCCAGTGGCTCCGAGACACT	3360
Db	3703	AGTGTGCTGTGATGGCAAAAGATTAAGCAACCTTAATGTCCAGTGGCTCCGAGACACT	3762
OY	3361	GGGCACTCGTGTCTCAGAGAGCCCATCTGTTGACAGCAGCACTTCGCGAGAACTTGGCTTA	3420
Db	3763	GGGCACTCGTGTCTCAGAGAGCCCATCTGTTGACAGCAGCACTTCGCGAGAACTTGGCTTA	3822
OY	3421	TGGAGACAACAGCCGAGGTCTGTATCCATGAAGAGATTATGCAAGCAGCCAAAGAGGGCAA	3480
Db	3823	TGGAGACAACAGCCGAGGTGTGTCTACAGAAAGAGATCTGTAGGGGAGCAAAAGAGGCCAA	3882
OY	3481	CATACACCACTTCATGCAGAACACTCCCTGAGAAATATCAACACACAGATAGAGACAAAG	3540
Db	3883	CATACATGCCCTTCATGCAGAACACTGTCTATTAATATATACACTAAAGTAGAGACAAAG	3942
OY	3541	AATCCAGCTCTCTGGTGGGCGACAAAGCGCATTTGCCATAGCTGGCGCTCTTTGTAGCA	3600
Db	3943	AATCTAGCTCTCTGGTGGGCGACAAAGCGCATTTGCCATAGCTGGCGCTCTTTGTAGCA	4002
OY	3601	GGCTCATATTTTCTTTTGGATGCAAGCTTCATCAGCTCTGGATTAAGAAAGTGAAGAAAGT	3660
Db	4003	GGCTCATATTTTCTTTTGGATGCAAGCTTCATCAGCTCTGGATTAAGAAAGTGAAGAAAGT	4062
OY	3661	TGTCCAGAGAGCCCTGGACAAAGCCAGAGAAAGGCGCACCTGCATTTGTGATGGCCACCG	3720
Db	4063	TGTCCAGAGAGCCCTGGACAAAGCCAGAGAAAGGCGCACCTGCATTTGTGATGGCTCACCG	4122
OY	3721	CTTGTCACCATTCAGAAATGCAATTTAATTAATTAATTAATTAATTAATTAATTTTAAAT	3780
Db	4123	CTTGTCACCATTCAGAAATGCAATTTAATTAATTAATTAATTAATTAATTTTAAAT	4182
OY	3781	GCATGGGCACATCAACAGGTGCTGGGCCCAAGAAAGGCATCTATTTTCCATGTCAGTGT	3840
Db	4183	GCATGGGCACATCAACAGGTGCTGGGCCCAAGAAAGGCATCTATTTTCCATGTCAGTGT	4242
OY	3841	CCAGGCTGAGCAAAAGCGCTAGTAAGTGTGGCCATCTAGAGCTTAATTAATTTTAAAT	3900
Db	4243	CCAGGCTGAGCAAAAGCGCGCTAGTAAGTGTGGCCATCTAGAGCTTAATTAATTTTAAAT	4302
OY	3901	AATTGTGTTAAACATGGCATTTATCAAAAGTTTAAAGGTGAGACTTACTGGAAGAACT	3960
Db	4303	AATT--TGTTTAAATATGACATTATTCOAAGTTAAA-----AGCAAAACACTTACAGATT	4355
OY	3961	ATGTAGAACACTACCTGTTTAAACATTTCTTGCTGCACAACGAGAGATCATCCACAAAGTTCAG	4020
Db	4356	ATGTAGAACAGTATCTGTTTAAACATTTCCCTCAGTCAAGTTCAGAGCTTTC-----	4403
OY	4021	AATCTTCAGATTTTAATTAATTAAGAACCAA--AAGAAACATTTATCGATGAGAAATAAT	4078
Db	4404	-----AGAACCTTCGTAATTAATTAAGAACCAAGAGTGAACATCAAGTGAAGAGAAAT	4458
OY	4079	ACTGTGTTAATTTGATTAATAAATTAATTAAGACTAATTCAAGTCAATTTGTATAATAT	4138
Db	4459	CATAGTTTAACTGCATTTAATTTTATTAACAGAAATTAAGATTTAATTAAGATTAAT	4518
OY	4139	TGTAATAATTTTGTATAATTTT-----ATTGTAACTTACTGCTTGGCTGAAGAT	4190

D	db	4519	ATGTGTAATTGGTTTAAATTTTCCTCCCATTTTGCGACTGTAAGTACGTACCCTTGCTTAAGAAGAT	4578
OY	4191	TATAGAAAGTGCAATAAAGTACTG- AATGTTGAATAAAGTGTAGTATATAATAAACAA	4249	
D	db	4579	TATAGAAAGTAGCAAAAAGATATGAATGTGTGCATTAAGT--TCTATATAATAACTPA	4635
OY	4250	ACTTTTATATCAAAAAAAAAAAAAAA 4279		
D	db	4636	ACTTTCATGTGAAAAAAAAAAAAAAAA 4665	
R	RESULT	10		
E	E02326		4378 bp	RNA linear PAT 29-SEP-1997
L	LOCUS	E02326		
D	DEFINITION	Multidrug resistance relating gene derived from human normal cells.		
A	ACCESSION	E02326.1 GI:2170561		
V	VERSION	JP 1990100680-A/1.		
K	KEYWORDS	Homo sapiens.		
S	SOURCE	Homo sapiens.		
O	ORGANISM	Homo sapiens. Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.		
R	REFERENCE	Ueda,K. and Komano,T. HUMAN NORMAL CELL-DERIVED MDR RELATED GENE Patent: JP 1990100680-A 1 12-Apr-1990;		
A	AUTHORS	JOURNAL		
C	COMMENT	SHUTOHY LTD OS Homo sapiens PN JP 1990100680-A/1 PD 12-Apr-1990 PF 05-Oct-1988 JP 1986251475 PI UEDA KAZUMITSU, KOMANO TORU PC C12N15/12,C12N1/21,C12O1/68; CC strandedness: Single; CC topology: Linear; CC *source: tissue.type=Adrenal gland; CC *source: clone=SM1132; FH key Location/Qualifiers FE FI 5'UTR 1..137 FT CDS 138..3980 FT /gene='Multidrug resistance relating gene' FT mat.peptide 138..3977 FT /gene='Multidrug resistance relating gene' FT FT FT /gene='MDRAL'		
F	FEATURES	FT 3'UTR Location/Qualifiers source 1..4378 /organism='Homo sapiens' /db_xref='taxon:9606' BASE COUNT 1315 a 818 c 1057 g 1188 t ORIGIN		
Q	Query Match	78.1%; Score 3341.2; DB 6; Length 4378;		
B	Best Local Similarity	87.6%; Pred. No. 0; Mismatches 488; Indels 44; Gaps		
M	Matches 3758; Conservative	0;		
OY	1	GGAACGGAGAGTCGGGATGATCTCGAAGAGGCCGTPAAGGGAGTCAGAGAAACAATT 60		
D	db	122 GGACCGGAGAGTCGGATGATCTTGAAGGGAGCCGAATGAGAGCAAGAAGAA 181		
OY	61	CTGGAANAATGGCGAAAAAAGTAAAAAAAATGAGAAAGAAAGAAACCAACTGCGAG 120		
D	db	182 CTTTTTAAACGTGAACAATTAAGAAGGAAAAAATTAAGAGGAAAAAGCAACTGTCAG 241		
OY	121	CACATTGGCAATGTTTCGCTATTCAAATAGGCTGATAGTGGTTGATATGTTGGGGAC 180		
D	db	242 TGTATTTTCAATGTTTCGCTATTCAAATAGGCTGATAGTGGTTGATATGTTGGGGAC 301		
OY	181	AATGGCGCATCATTCATGAGAGCTCACCTCTCATGATGCTGGTTTGGAAACAT 240		
D	db	302 TTGGCTGCCATCATTCATGAGAGCTGCAATCTCTCATGATGCTGGTTGGAGAAAT 361		

OY	241	GAGAGATACCTTGGCAATGACAGAAATTTCAGAGAAACAAACTTTCGATTAATAATTA	300
Db	362	GACAGATATCTTTGGCAATGACAGAAATTTAACAATCTGATGTG-----CAACATCAC	415
OY	301	TGAAGATTAATGCAACAATACACAACTTTCATCAACCATCTGAGAGAGAAATGACAC	360
Db	416	TAAATAAAGTATATCATATATACAGAGGTTCTTTCATGTAGATCTGGAGAGACATACAG	475
OY	361	GTATGCCATTAATTAACAGTGGATCGSTCTCGCATCGCTGTGGTCTCTTAATCCAGAT	420
Db	476	GTATGCCATTAATTAACAGTGGAAATGTGCTCGTGGTGTCTGTTGCTTACATTCAGAT	535
OY	421	TTCATCTGAGGCCCGGACGACAGAGAAACAGATATCTCAAAATTATGAAACAATTTTCCA	480
Db	536	TTCATTTATGTGTCCTGGAGAGCTGGAAAGACAAATACACAAATTAAGAAACAGTTTTTCA	595
OY	481	TGCTATCATGCGACAGAGATTTGCTGTGGTTGACSTGCATGACCTTGGGAGCTTAACAC	540
Db	596	TGCTATAAAGCGACAGAGAGATAGCGCTGGTTTATGTGCACATTTGGGAGCTTAACAC	655
OY	541	CCGGGTCACAGACAGATGTGCTCCAAAATCATTAAGAAATTTGGCAGCAAAATTTGAAATGT	600
Db	656	CCGACTTACAGATGATGTCTCCAAAGATTAATGAAGAAATTTGTGACAAAATTTGAATGT	715
OY	601	CTTTCACATCAAGCAACATTTTTCACCGGTTTATGATGGGGTTTACACGTGGTTTGGAA	660
Db	716	CTTTCAGTCAATGAAGCAACATTTTTCACGTGGTTTATGATAGAAATTTACGTGGTTTGGAA	775
OY	661	GCTAACCCCTGTGATTTTGGCATACACCCCTGTCTTGGACTTTCACGCGCATCTGGGC	720
Db	776	GCTAACCCCTGTGATTTTGGCATACAGTCCCTGTCTTGGACTTTCACGCTGTCTGGGC	835
OY	721	TAAGATACATCTTCATTTACTGATTAAGAAACCTCTGCGCATGCAAAACCTGAGCAGT	780
Db	836	TAAGATACATCTTCATTTACTGATTAAGAAACCTTACGATGCAAAACCTGAGCAGT	895
OY	781	AGCTGAAGAAGCTTAGAGCAATACAGACTGAGACTGTGATGCTTGGAGAGCAAAAGAAAGA	840
Db	896	AGCTGAAGAAGCTTGGCAGCAATTAAGACTGTGATGCTTGGAGAGCAAAAGAAAGA	955
OY	841	ACTTGAAGGTACACAACAAATTTAGAGAGCTTAAGAGATTTGGATTAAGAAAGCTAT	900
Db	956	ACTTGAAGGTACACAACAAATTTAAGAGAGCTTAAGAGATTTGGATTAAGAAAGCTAT	1015
OY	901	CACGGCCAACTTTCATTTATGGTGCCCTTTCATTTGATGATGCAATCATATGCTGTGC	960
Db	1016	TACAGCAATATTTCTATAGAGTGCTGTTCTCTGCTGATCATGATCATCTATGCTGTGC	1075
OY	961	TTTCTGATTTGGAGACCTCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG	1020
Db	1076	CTTCTGATTTGGAGACCTCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG	1135
OY	1021	TGCTCTCTTCTCTGATTAATTTGGGCGCTTATGATTTGGACAGGCAATCCCAAGCATTA	1080
Db	1136	TGATATCTTCTTCTGATTAATTTGGGCGCTTATGATTTGGACAGGCAATCCCAAGCATTA	1195
OY	1081	AGATTTGCAAAACGAGAGAGAGACGCTTATGAATCTTCAAGATTAATGACATTAAC	1140
Db	1196	AGCATTTGCAAAATGCAAGAGAGAGACGCTTATGAATCTTCAAGATTAATGATTAATGAC	1255
OY	1141	AAACATTTGACACCTTTCGAAAGAGTGGACATTAACCAATATTTTAAGGAAATTTGGA	1200
Db	1256	AAATATTTGACACCTTTCGAAAGAGTGGACCAACCAATATATTTTAAGGAAATTTGGA	1315
OY	1201	ATTCAAAAATGTTCACTTCAGTTACCCCTTCTGAAAAGAGATTAAGATCTTAAGGGCT	1260
Db	1316	ATTCAAAAATGTTCACTTCAGTTACCCATCTCGAAAAGAGATTAAGATCTTGAAGGCT	1375
OY	1261	CAACCTGAAGGTTCAGAGTGGGCGAGACAGTGTGCTGTTGGAGACAGTGGCTGGGGA	1320
Db	1376	GAACCTGAAGGTTCAGAGTGGGCGAGACAGTGTGCTGTTGGAGACAGTGGCTGGGGA	1435

QY	1321	GACACAGACCCTGCAGCTGATGACAGAGGCGCTATACCCCAACAAATGCGATCGTAT	1380
Db	1436	GAGCACAACAGTCCAGCTGATGACAGAGGCTCTATACCCCAACAGAGGAGTGCAGTGT	1495
QY	1381	TGATGGACAGGACATTTAGGACCAATAATGAAGGCATCTTCGGAAATTACTGTGTGT	1440
Db	1496	TGATGGACAGGATATTATAGACCAATAATGAAGGATTCTACGGAAATCTTGTGTGTGT	1555
QY	1441	GAGTCAGAGAGCCTGTGTGTTTGGCAACAGCAATACCTGAATAACCTTGTGCTAGGCGCGGA	1500
Db	1556	GAGTCAGAGAACCTGTATTGTTTGGCAACAGCAATACCTGAATAACCTTGTGCTAGGCGCGGA	1615
QY	1501	AAATGTCCACCATGGATGAGATTTGAGAAAGCTGTAAAGAACCCAACTCCATGATTTAT	1560
Db	1616	AAATGTCCACCATGGATGAGATTTGAGAAAGCTGTCAAGAAACCCAACTCCATGATTTAT	1675
QY	1561	CATGAACCTACCTAATTAATTTGACACTGCTGTTGAGAGAGAGGGCCACGCTGAGTGG	1620
Db	1676	CATGAACCTCCCTCAATTAATTTGACACCCCTGGTTGAGAGAGAGGGCCACGCTGAGTGG	1735
QY	1621	TGCAGAAACAGAAATTCGCAATTCGTGCGGCCCTGGTGTGCACACCCCAAGATTCCTCT	1680
Db	1736	TGGCGAAGAACAGAGGATTCGCATTCAGAGTGCCCTGGTGTGTAAACCCCAAGATTCCTCT	1795
QY	1681	GCTGTGATGAGGACAGCTAGCTCTGACACCTGAAGTGAACAGTGGTGTAGGTGGCCT	1740
Db	1796	GCTGTGATGAGGACAGCTAGCTCTGACACCTGAAGTGAACAGTGGTGTAGGTGGCCT	1855
QY	1741	GGATTAAGGCCAGAAAAAGGCCGACATCACTTGATGTAGCTCATGTTTGTCTACAGTTG	1800
Db	1856	GGATTAAGGCCACAAAAAGGTCGGACACACATTCGATGTAGCTCATGTTTGTCTACAGTTG	1915
QY	1801	TAAATGCCATGTCAATTCGTGTGTTTGTATGTAGAGAGCATTTGTGAGAAAGAAATCATGA	1860
Db	1916	TAAATGCCATGTCAATTCGTGTGTTTGTATGTAGAGAGCATTTGTGAGAAAGAAATCATGA	1975
QY	1861	TGAATCATGAAGAGAGAGGCACTTACTTCAAACTTGTCAATTCGACACAGAGAGAA	1920
Db	1976	TGAATCATGAAGAGAGAGGCACTTACTTCAAACTTGTCAAACTTGTCAAACTGACACAGAGAA	2035
QY	1921	TGAATTTGAGTTAGAAAAATTCGCACCTGTGTGAATCCAAAAAGTAAAGTATCCCTTGAAT	1980
Db	2036	TGAATTTGAGTTAGAAAAATTCGCACCTGTGTGAATCCAAAAAGTAAATGATCCCTTGAAT	2095
QY	1981	GTCGCCAAAAAGTTGAGGTCCTCACTTAAATAAAAAGAAAGCACTGCGAGGAGTTACA	2040
Db	2096	GTCGCCAAAAAGTTGAGGTCCTCACTTAAATAAAAAGAAAGCACTGTCAGAGAGTCTCG	2155
QY	2041	TGCACCAAGAGCCAGACAGACAGAGCTTGTGTACAAAAAGAGCACTTGATGAGATGTAC	2100
Db	2156	TGCATCCAAAGCCCAAGACAGACAGAGCTTGTGTACAAAAAGAGCTTGTGTAAAGTATAC	2215
QY	2101	TCCAGTTTCCTCTCTGAGAGATTCGAGAGCTGAAGCTCAACTAATAGGCGCTAATTTGTGT	2160
Db	2216	TCCAGTTTCCTCTCTGAGAGATTCGAGAGCTGAAGCTAATTACTAATAGGCGCTAATTTGTGT	2275
QY	2161	TGGTAAATTTTGGCTATTATAAAGGAGAGGCCGCAACACACATTTCAATTAATTTTC	2220
Db	2276	TGGTAAATTTTGGCTATTATAAAGGAGAGGCCGCAACACACATTTCAATTAATTTTC	2335
QY	2281	CAGTGTTCGTAATGTTTGTATAGTCCCTTGGAAATTAATTTCTTTTATTCATTTTCTCCA	2340
Db	2396	CAGTGTTCGTAATGTTTGTATAGTCCCTTGGAAATTAATTTCTTTTATTCATTTTCTCCA	2455
QY	2341	GGGCTTCACATTTGGCAAAAGCTGGGAGATCCCTCACTAAGGCGCTTGATACATGGTTTT	2400
Db	2456	GGGCTTCACATTTGGCAAAAGCTGGGAGATCCCTCACTAAGGCGCTCGATATATGGTTTT	2515
QY	2401	CAGTCCATGCTGACACAGATGTCAAGTGTGTGATGACCTTAATAAACCACTGAGAC	2460

||||| 2575
CGATCCATGCTGACAGCAGGATGTGAGTGGTTGATGAGCCCTAAAAACACCACCTGAGAC 2575
QY 2461 ATTGACACCAAGGCTTGGCAATGATGCGGCTCAAGTTAAAGGGGCTATGTTCCAGGCT 2520
||||| 2635
2576 ATTGACTACAGGCTCGCCCAATGATGCTGCTCAAGTTAAAGGGGCTATGTTCCAGGCT 2635
QY 2521 TGTGCTATTACCCGAAATATAGCAAACTTGGGACAGCATTTATATCTTAACTTA 2580
||||| 2695
2636 TGTGTAATTTACCGAAATATAGCAAACTTGGGACAGCAATATATATCTTCACTTA 2695
QY 2581 TGGTGGCAATTAACACTTTTACTCTTAGCAATTTGATACCATCTTTCATTAACAGAGT 2640
||||| 2755
2696 TGGTGGCAATTAACACTTTTACTCTTAGCAATTTGATACCATCTTTCATTAACAGAGT 2755
QY 2641 TGTGAAATGAAATGTTCTGAGCAAGCACTGAAAGATAGAAAGAGCTAGAAAGAGC 2700
||||| 2815
2756 TGTGAAATGAAATGTTCTGAGCAAGCACTGAAAGATAGAAAGAGCTAGAAAGAGTTC 2815
QY 2701 TGGGAAGATTGCTACAGAAAGCAATGCAAACTTCCGAACCTTGTGTTCTTTGACTCGGA 2760
||||| 2875
2816 TGGGAAGATTGCTACAGAAAGCAATGCAAACTTCCGAACCTTGTGTTCTTTGACTCAGGA 2875
QY 2761 GCAGAAATTTGAATACATGATGACAGAGTTTGGCAAGTACCATACAGAACTCTTTGAG 2820
||||| 2935
2876 GCAGAAATTTGAATACATGATGATGCTCAGAGTTTGGCAAGTACCATACAGAACTCTTTGAG 2935
QY 2821 GAAACACACATCTTGGGGTCTCATTTTCTATCACCCGAGCAATGATGATTTTCTTA 2880
||||| 2995
2936 GAAACACACATCTTGGGAATTAATTTCTTCCACCGCAAGCAATGATGATTTTCTTA 2995
QY 2881 TGTGCGCTTTCGGGTTTGGTGGCTACTTGTGTGCAAAATGAGTTCATGAACTTTCAAGA 2940
||||| 3055
2936 TGTGCGCTTTCGGGTTTGGTGGCTACTTGTGTGCAAAATGAGTTCATGAACTTTCAAGA 3055
QY 2941 TGTGCTTTTGGATTTCTCAGTATGTTCTTTGCTTGGCCATGCGAGTGGGACAGTCTTC 3000
||||| 3115
3056 TGTGCTTTTGGATTTCTCAGTATGTTCTTTGCTTGGCCATGCGAGTGGGACAGTCTTC 3115
QY 3001 ATTTGCTCTGACATTTGCAAGGCAAAAGTATCAGAGCCAGCTCATCATGATCTTGA 3060
||||| 3175
3116 ATTTGCTCTGACATTTGCAAGGCAAAAGTATCAGAGCCAGCTCATCATGATCTTGA 3175
QY 3061 AAAAAGCCCTGATTTGACAGCTACAGCCCTCAGAGGCTCAAGCAAAATCGTTGAAG 3120
||||| 3235
3176 AAAAAGCCCTGATTTGACAGCTACAGCCGAGGCTTAATGCGCAACATTTGAAG 3235
QY 3121 AAATGCAATTTAATGAGTGTGTTCACTATCCACTGACAGCAGATCCCGTGT 3180
||||| 3295
3236 AAATGCAATTTGAGTGTGTTCACTATCCACTGACAGCAGATCCCGTGT 3295
QY 3181 CCAAGGGCTGAGGCTGAGGTGAGAGAGGCGCAGAGCGTGGCCCTCGAGTACAGTGG 3240
||||| 3355
3296 TCAAGGAGCTAGGCTGAGGTGAGAGAGGCGCAGAGCGTGGCCCTCGAGTACAGTGG 3355
QY 3241 CTGAGGAAAGAGCAGTTGTTAGTCTCTAGAGCGCTTATAGACCCCTTGGCTGCTC 3300
||||| 3415
3356 CTGAGGAAAGAGCAGTTGTTAGTCTCTAGAGCGTTCACAGCCCTTGGAGAGGAA 3415
QY 3301 AGTCTAATTTGATGAGCAAGATTAAGCACTGAATGTCAGTGGCTCCGAGCAGACT 3360
||||| 3475
3416 AGTCTGCTTATGAGCAAGATTAAGCACTGAATGTCAGTGGCTCCGAGCAGACT 3475
QY 3361 GGGCATCTGCTCAAGAGCCCATCCGTTTGAAGCTGACAGATTTGCGAAGATTTGCTTA 3420
||||| 3535
3476 GGGCATCTGCTCAAGAGCCCATCCGTTTGAAGCTGACAGATTTGCGAAGATTTGCTTA 3535
QY 3421 TGGAGACAAGAGCGGCTGATCATGAGAGATTTATGAGAGCCCAAGAGAGGCCAA 3480
||||| 3595
3536 TGGAGACAAGAGCGGCTGATCATGAGAGATTTATGAGAGCCCAAGAGAGGCCAA 3595
QY 3481 CATACACACATTTATGAGAGCTCCCTGAAATATCAACACAGAGTAGAGACAAAG 3540
|||||

Db 3596 CATACATGCTTCATGAGTCACTGCTCCTAATATATAGCACTAAAGTAGAGACAAAG 3655
QY 3541 AACCCAGCTCTGCTGGGCGCAAGAAACAGCGATTTGCCATAGCTGGGCTTGTGAGA 3600
||||| 3715
3656 AACCTAGCTCTGCTGGGCGCAAGAAACAGCGATTTGCCATAGCTGGGCTTGTGAGA 3715
QY 3601 GCGCTATTTTGGCTTTTGGATGAGTCAATCAGCTCTGATGATCAGAAAGTAAAGGT 3660
||||| 3775
3716 GCGCTATTTTGGCTTTTGGATGAGTCAATCAGCTCTGATGATCAGAAAGTAAAGGT 3775
QY 3661 TGTCCAAAGAGCCCTGAGCAAGCCAGAGAGGCGGCACTGCTCATTTGATGAGTCAAGGA 3720
||||| 3835
3776 TGTCCAAAGAGCCCTGAGCAAGCCAGAGAGGCGGCACTGCTCATTTGATGAGTCAAGGA 3835
QY 3721 CTGCTCACCATCCAGAAATGCAATTTATGATGATGATGATGATGATGATGATGATGAT 3780
||||| 3895
3836 CTGCTCACCATCCAGAAATGCAATTTATGATGATGATGATGATGATGATGATGATGATGAT 3895
QY 3781 GCATGCAACATCAACAGCTGCTGCGCCAGAAAGGATCTATTTTCCATGAGTCAAGTGT 3840
||||| 3955
3896 GCATGCAACATCAACAGCTGCTGCGCCAGAAAGGATCTATTTTCCATGAGTCAAGTGT 3955
QY 3841 CGAGCTGAGCAAGAGCGCTAGTGAAGTGGCCATATGAGCTGTTAAATTTTAAAT 3900
||||| 4015
3956 CGAGCTGAGCAAGAGCGCTAGTGAAGTGGCCATATGAGCTGTTAAATTTTAAAT 4015
QY 3901 ATTTGTTTAAAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 3960
||||| 4068
4016 ATTTGTTTAAAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 4068
QY 3961 ATTTGTTTAAAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 4020
||||| 4116
4069 ATTTGTTTAAAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 4116
QY 4021 AGTCTTCAATTTTAAATTTTAAAGATGATGATGATGATGATGATGATGATGATGAT 4078
||||| 4171
4117 AGTCTTCAATTTTAAATTTTAAAGATGATGATGATGATGATGATGATGATGATGATGAT 4171
QY 4079 ACTGCTGATTTTAAATTTTAAAGATGATGATGATGATGATGATGATGATGATGAT 4138
||||| 4231
4172 ACTGCTGATTTTAAATTTTAAAGATGATGATGATGATGATGATGATGATGATGATGAT 4231
QY 4139 TGTATATTTTAAATTTTAAATTTTAAAGATGATGATGATGATGATGATGATGATGAT 4190
||||| 4291
4232 ATGCTATTTTAAATTTTAAATTTTAAAGATGATGATGATGATGATGATGATGATGAT 4291
QY 4191 TATGAGAGTGTAAAGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 4249
||||| 4348
4292 TATGAGAGTGTAAAGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 4348
QY 4250 ACTTTTATCAAAAAAATTTTAAAGATGATGATGATGATGATGATGATGATGATGAT 4279
||||| 4378
4349 ACTTTTATCAAAAAAATTTTAAAGATGATGATGATGATGATGATGATGATGATGATGAT 4378
Db

RESULT 11
AR091275 4669 bp DNA linear PAT 07-SEP-2000
LOCUS AR091275
DEFINITION Sequence 1 from patent US 5994088.
ACCESSION AR091275
VERSION AR091275.1 GI:10018030
KEYWORDS
ORGANISM Unknown.
SOURCE Unknown.
REFERENCE Unclassified.
1 (bases 1 to 4669)
AUTHORS Meheiner,E. and Roninson,I.B.
TITLES Methods and reagents for preparing and using immunological agents
specific for P-glycoprotein
JOURNAL Patent: US 5994088-A1 30-NOV-1999;
FEATURES Location/Qualifiers
1..4669
source /organism="unknown"

BASE COUNT	1393 a	894 c	1130 g	1252 t
ORIGIN				
Query Match	77.9% Score 3333.2; DB 6; Length 4669;			
Best Local Similarity	87.5%; Pired. No. 0;			
Matches 3753; Conservative	0; Mismatches 493; Indels 44; Gaps 8;			
OY	1 GGAGCGGAGGTGCGGATGATGATCCTGTAAGAGAGCGCGTAAGGGAGGAGGAGAGAACTT 60			
Db	409 GGAGCGGAGGTGCGGATGATGATCCTGTAAGGGAGGAGGAGAGAACTT 60			
OY	61 CTGGAATATGGGCAAAAAAGTAAAAAATGAGAAAGAAAGAAAGAAAGAAAGAAAGAA 468			
Db	469 CTTTAACTGACACATATAAGATGAAAAAGATGAAAAAGAAAGAAAGAAAGAAAGAA 528			
OY	121 CAGCTTGAATGTTTGGTATCAATCAATGGCTTGAATGATGATGATGATGATGATGAT 180			
Db	529 TGTATTTCAATGTTTGGTATCAATCAATGGCTTGAATGATGATGATGATGATGATGAT 588			
OY	181 AATGGCGCATATCATCATGAGCTGACCTCCCTCATGATGATGATGATGATGATGATGAT 240			
Db	589 TTTGGCTGCTATCATCATGAGCTGACCTCCCTCATGATGATGATGATGATGATGATGAT 648			
OY	241 GACAGATAGCTTTCGAATGAGAGATTCGAAGAAAGAAAGAAAGAAAGAAAGAAAGAA 300			
Db	649 GACAGATATCTTGGAAATGAGAGAAATTTGAAGATCTGATGATGATGATGATGATGATGAT 702			
OY	301 TGAAGATATACGAACATATACCAACATTCATCAACATTCATGAGAGAGAAATGACAC 360			
Db	703 TAAATGAAGTATATCAATGATACAGAGGTTCTCATGATGATGAGAGAGAAATGACAC 762			
OY	361 GATAGCTATATATACAGTATGAGTATGAGTATGAGTATGAGTATGAGTATGAGTATGAG 420			
Db	763 GATAGCTATATATACAGTATGAGTATGAGTATGAGTATGAGTATGAGTATGAGTATGAG 822			
OY	421 TTGATTCGAGTCTGCGGAG 480			
Db	823 TTGATTCGAGTCTGCGGAG 882			
OY	481 TGCATATCAG 540			
Db	883 TGCATATCAG 942			
OY	541 CCGGCTCAG 600			
Db	943 CCGGCTCAG 1002			
OY	601 CTTTCACTCAATAGCAACATTTTTCACCGGTTTATAGTGGGTTTACAGCTGGTGGAA 660			
Db	1003 CTTTCACTCAATAGCAACATTTTTCACCGGTTTATAGTGGGTTTACAGCTGGTGGAA 1062			
OY	661 GCTAACCTTGTGATTTGGCCATGAGCCCTGTTCTTGAGCTTTCAGCCCGCATCTGGGC 720			
Db	1063 GCTAACCTTGTGATTTGGCCATGAGCCCTGTTCTTGAGCTTTCAGCCCGCATCTGGGC 1122			
OY	721 AAGATATATATCTTCAATTTAGTAAAGAACTTGGCTATGCAAAAGCTGGAGAGT 780			
Db	1123 AAGATATATATCTTCAATTTAGTAAAGAACTTGGCTATGCAAAAGCTGGAGAGT 1182			
OY	781 AGCTGAAGAGTCTTAGCAACATGAGAACTGATGATGATGATGATGATGATGATGATGAT 840			
Db	1183 AGCTGAAGAGTCTTAGCAACATGAGAACTGATGATGATGATGATGATGATGATGATGAT 1242			
OY	841 ACTTGAAGAGTCTTAGCAACATTTTGAAGAAAGCTTAAAGAAAGAAAGAAAGAAAGCTAT 900			
Db	1243 ACTTGAAGAGTCTTAGCAACATTTTGAAGAAAGCTTAAAGAAAGAAAGAAAGAAAGCTAT 1302			
OY	901 CAGGCGCAACATTTCTATTTGGTCCCTTCTTATGATATGATGATGATGATGATGATGATG 960			
Db	1303 TACACCGCAATTTCTATTTGGTCCCTTCTTATGATATGATGATGATGATGATGATGATG 1362			
OY	961 TTTCTGATAGGAGCTTCTGCTCTCAGTGAATATCTTATGAGCAAGTACTGAC 1020			
Db	1363 CTTCTGATAGGAGCTTCTGCTCTCAGTGAATATCTTATGAGCAAGTACTGAC 1422			
OY	1021 TGTCTTCTTCTTCTGATTAATTTGGGCTTTTATGATGAGAGAGAGAGAGAGAGAGAGAG 1080			
Db	1423 TGTCTTCTTCTTCTGATTAATTTGGGCTTTTATGATGAGAGAGAGAGAGAGAGAGAGAG 1482			
OY	1081 AACATTTGCAAG 1140			
Db	1483 AACATTTGCAAG 1542			
OY	1141 AAGCATGACAGCTATTCGAG 1200			
Db	1543 AAGCATGACAGCTATTCGAG 1602			
OY	1201 ATTCGAAATGTTTCACTTCACTTACCTTCTGCAAAAGAGTTAAGATCTTAAAGGCTT 1260			
Db	1603 ATTCGAAATGTTTCACTTCACTTACCTTCTGCAAAAGAGTTAAGATCTTAAAGGCTT 1662			
OY	1261 CAACCTGAAGGTTGAG 1320			
Db	1663 CAACCTGAAGGTTGAG 1722			
OY	1321 GAGCAGACCTGTCAGCTGATGAGAGAGGCTTATGACCCAGATGGCATGCTGTAT 1380			
Db	1723 GAGCAGACCTGTCAGCTGATGAGAGAGGCTTATGACCCAGATGGCATGCTGTAT 1782			
OY	1381 TGAATGACAGAGATATGAG 1440			
Db	1783 TGAATGACAGAGATATGAG 1842			
OY	1441 GAGTCAGAGAGCTGTGTGTTTGGACACAGATAGCTGAAAGAAATTCGATGAGT 1500			
Db	1843 GAGTCAGAGAGCTGTGTGTTTGGACACAGATAGCTGAAAGAAATTCGATGAGT 1902			
OY	1501 AATATGACAGAGATATGAG 1560			
Db	1903 AATATGACAGAGATATGAG 1962			
OY	1561 CAGTGAACCTGCTATTAATTTGACACTGCTGTTGAGAGAGAGAGAGAGAGAGAGAGAG 1620			
Db	1963 CAGTGAACCTGCTATTAATTTGACACTGCTGTTGAGAGAGAGAGAGAGAGAGAGAGAG 2022			
OY	1621 TGGAG 1680			
Db	2023 TGGAG 2082			
OY	1681 GCTGATGAG 1740			
Db	2083 GCTGATGAG 2142			
OY	1741 GATTAAG 1800			
Db	2143 GATTAAG 2202			
OY	1801 TAAATGCGAGTCAATGCTGTTTATGATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1860			
Db	2203 TAAATGCGAGTCAATGCTGTTTATGATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2262			
OY	1861 TGAATCATGAG 1920			
Db	2263 TGAATCATGAG 2322			
OY	1921 TGAATTTGAGTGAAGAAATGCACTGATGATGATGATGATGATGATGATGATGATGATG 1980			
Db	2323 TGAATTTGAGTGAAGAAATGCACTGATGATGATGATGATGATGATGATGATGATGATG 2382			
OY	1981 GTCCTCAAAAG 2040			
Db	2383 GTCCTCAAAAG 2442			
OY	2041 TGCACCAAG 2100			
Db	2443 TGCACCAAG 2502			

QY	2101	TCGAGTTTCGCTGCGGAGATCTCTGAAGCTGAACCTCACTGATGGCTTATTTTGTCG	2160
Db	2503	TCGAGTTTCGCTTGGAGGATATGAAGCTAAATTTAACTGAATGGCTTATTTTGTCG	2562
QY	2161	TGGTATATTTTGTCTATTTATAAGGAGGCCGGAACGAGATTTTCAATATATTTTC	2220
Db	2563	TGGTATATTTTGTCTCAATTAATAATGAGGCTTCGACACGACATTTTGCATATATTTTC	2622
QY	2221	AAGGATTTATGAGATCTTTTACCCAGATGAGATCTCTGAACAAACGACAAATATGTA	2280
Db	2623	AAAGATTTATAGGGTTTTTACAAGAATATGATATCTCGAAACAAACGACAAATATGTA	2682
QY	2281	CATGTTTCTGTATTTGTTCTTAGTCCTGGATATTTCTTTATATACATTTTCTCTCA	2340
Db	2683	CTTGTTTTCACTATTTGTTTCTTAGCCCTTGGATATTTCTTTTATACATTTTCTCTCA	2742
QY	2341	GGGCTTCACTTTGGCAAGCTGGGGAGATCTCTACTAAGCGCTTCGATACATTTGTTTT	2400
Db	2743	GGGTTTCACTTTTGGCAAGCTGGGAGATCTCTACCAAGCGCTTCGATATATTTGTTTT	2802
QY	2401	CAGATCCATGCTGAGAACGAGATGTAGCTGGTTTGATGACCTTAAACACCACTGGAGC	2460
Db	2803	CCGATCCATGCTGAGACGAGATGTAGTTGGTTTCATGACCTTAAACACCACTGGAGC	2862
QY	2461	ATTGACAAACGAGCTTGGCAATGATGCGCTCAAGTTAAAGGGCTTATGGTTCCAGGCT	2520
Db	2863	ATTGACTACAGGCTTGGCAATGATGCTCAAGTTAAAGGGCTTATAGTTCCAGGCT	2922
QY	2521	TGCTGTCATTAACCCAGATATATGCAAAATCTTGGGAGGCAATTAATATCTTAATCTA	2580
Db	2923	TGCTGTCATTAACCCAGATATATGCAAAATCTTGGGAGGCAATTAATATCTTCACTTA	2982
QY	2581	TGTTGGCAATTAACATTTTACTCTTACCAATTTGACCAATCATTTGCAATATGACGAGT	2640
Db	2983	TGTTTGGCAATTAACATTTTACTCTTACCAATTTGACCAATCATTTGCAATATGACGAGT	3042
QY	2641	TGTTGAATGAATATGTTTGTCTGGACAGACCTGAAGATTAAGAAAGCTAAGAAAGGAGC	2700
Db	3043	TGTTGAATGAATATGTTTGTCTGGACAGACCTGAAGATTAAGAAAGCTAAGAAAGGAGC	3102
QY	2701	TGGGAAATTTCTTCAAGAACGATCGAAACCTTCCAACTGTGTTTTCTTTGACTCGGA	2760
Db	3103	TGGGAAATTTCTTCAAGAACGATCGAAACCTTCCAACTGTGTTTTCTTTGACTCGGA	3162
QY	2761	GCAGAAATTTGAATACATGTAATGCAAGACTTTGCAAGTACCAATAGAAACCTTTGAG	2820
Db	3163	GCAGAAATTTGAATACATGTAATGCTCGAGATTTGCAAGTACCAATAGAAACCTTTGAG	3222
QY	2821	GAAACACACATCTTGGGGGCTCATTTTGTATACACGACGACGACANTGATTTTCTTA	2880
Db	3223	GAAACACACATCTTGGGAATTAACATTTTCTTCCACCGACGACATGATGATTTTCTTA	3282
QY	2881	TGCTGAGCTTTTCCGGTTTGGTGCCTTACTTGTGTGGCAAAATGAGTCAATGAACCTTACGA	2940
Db	3283	TGCTGAGCTTTTCCGGTTTGGAGCCTACTTGTGTGGCAAAATGAGTCAATGAACCTTACGA	3342
QY	2941	TGCTCTTTTGTATTTCTAGTATTTGTCTTGTGTGGCAATGGCACTGGGGGAGCTCACTTC	3000
Db	3343	TGCTCTTTTGTATTTTCTAGCTTTGTCTTGTGTGGCAATGGCGCTGGGGGAGCACTTC	3402
QY	3001	ATTTTGTCTGACTATGCCAAAGCCAAAGTATACGACAGCCGACGTCATCATGATCATTTGA	3060
Db	3403	ATTTTGTCTGACTATGCCAAAGCCAAATATACAGACCCGACATCATGATCATTTGA	3462
QY	3061	AAAAAGCCCTCTATTTAGACAGCTACAGCCCTACGGCTCAAGCCAAATACGTTGGAAG	3120
Db	3463	AAAAAGCCCTTATTTAGACAGCTACAGCCGGAAGGCTTATGCGGAACACATTTGGAAG	3522
QY	3121	AAATGTGACATTTAATGAGGTGCTGTTTCAATATGCACTGACAGACATCCCGTGT	3180
Db	3523	AAATGTGACATTTTGAAGTTGATTTATTCACATTTCCACACCGGACGATCCGAGTGT	3582

OY	3181	CCAGGGCTGACCTTCAGAGTGAAGAGGCGCAGACGTGGCCCTGTAGTACGACATGG	3240
Db	3583	TCAGGAGCTGAGCCTTGGAGGTGAAGAGGCGCCAGACGCTGGCTCTGGTGGGACAGTGG	3642
OY	3241	CTGTGGGAAGAGCACAGTTGTTACGCTCCTAGAGGCGCTCTATGACCCCTTGGCTGGTTTC	3300
Db	3643	CTGTGGGAAGAGCACAGTGGTCCAGCTCTGGAGCGGTTCTACGACCCCTTGGCAGGAA	3702
OY	3301	AATGCTAATTTGTGGCAAGAGATTAAGACACCTGCAATGTCCAGTGGCTCCGACACACT	3360
Db	3703	AGTGGCTGTTATGGCAAAAGAAATTAAGCGCACTGAATGTTCACTGGCTCCGACACACT	3762
OY	3361	GGGCACTGCTGTCTCAGAGCCCATCTCTTTGACTGTGAGATTTGCCGAGAACTTGCCTA	3420
Db	3763	GGGCACTGCTGTCCAGAGGCCCATCTCTTTGACTGTGAGCACTTGCCTGAGAACTTGCCTA	3822
OY	3421	TGGAGCAACACGCGGGTGTATCCATGAAGAGATTATGACGACCCACAGAGGCTCAA	3480
Db	3823	TGGAGCAACACGCGGGTGTATCCAGGAAGAGATCTGTGGGCGACCAAGAGGCGCAA	3882
OY	3481	CATACACACCTTCATGAGACACTCCCTGAGAAATACAAACCCAGATGAGAGCAAAAG	3540
Db	3883	CATACATGCCCTTCATGAGACTCACTGCTAATTAATATGCACTTAAGTGGAGCAAAAG	3942
OY	3541	AACCCAGCTCTCTGGTGGGCGACAAACAGCGCACTTGCCTAGCTGCGGCTTGTTAACA	3600
Db	3943	AACCACTCTCTGGTGGGCGCAAAACAGCATTCGCCTAGCTGCGGCTTGTTAACA	4002
OY	3601	GCCCTCATATTTTCTTTGGATGACCTACATCAGCTCTGGATTACAGAAAGCAAAAGGT	3660
Db	4003	GCCCTCATATTTTCTTTGGATGACCCAGCTGACCTGGATTACAGAAAGTGAAGAGT	4062
OY	3661	TGTCCAAAGAGCCCTGGACAAAGCCAGAGAGGCGCACCTCATTTGTGATCGCCACCG	3720
Db	4063	TGTCCAAAGAGCCCTGGACAAAGCCAGAGAGGCGCACCTCATTTGTGATCGCCACCG	4122
OY	3721	CTGTCCACCATTCAGAAATGCAATTTAATAGTGGTTCAGAAATGGCAAGTCAAGGA	3780
Db	4123	CTGTCCACCATTCAGAAATGCAATTTAATAGTGGTTCAGAAATGGCAAGTCAAGGA	4182
OY	3781	GCATGGCACACATCAACAGCTGCTGGCGCCAGAAAGGCATATTTTCCATGGTCACTGT	3840
Db	4183	GCATGGCACACATCAACAGCTGCTGGCGCCAGAAAGGCATATTTTCCATGGTCACTGT	4242
OY	3841	CCAGGCTGGAGCAAGCGCTAGTGAAGCTGTGGCATTTGAGCGTTAAATATTTTAAAT	3900
Db	4243	CCAGGCTGGAGCAAGCGCGCAGTGAACCTGACTGTATGAGATSTTAAATACTTTTAAAT	4302
OY	3901	AATTTGTTAAACATGAGATTTATCAAGTATTAAGAGTGAAGCACTTACTGCAAAACT	3960
Db	4303	ATT- -TGTTAAATATGACATTTATTCAAACTTAA- - - - -AGCAAACTTACAGAAAT	4355
OY	3961	ATTGTAAGTACCTGTTTAAACATTTCTTGCTGCAACTGAAATCATTCGCCAAAGTTCAG	4020
Db	4356	ATTGTAAGTATCTGTTTAACTTTCTTCAGTCAAGTTCACAGTTC- - - - -	4403
OY	4021	AGTCTTCAGATTTTATTAATTAAGGACCAA- - - - -AAGAACATTTATCGATGAAATAAT	4078
Db	4404	-----AGAGCTTCGTAATTAAGGAGACAGAGTGAAGACATCATCAAGTGAAGAAAT	4458
OY	4079	ACTGGTGAATTTGCAATTTAAAAATTATAGACTAATTCAAAGTGAATTTGTTAAATAAT	4138
Db	4459	CATAGTTTAAACGCAATTTAAATTTTATACAGATTAAAGTAAATTTTAAAGATTA	4518
OY	4139	TGTATTAATTTTGTATTAATTT- - - - -ATTGTACTTACGCTTGTGCGAAGAT	4190
Db	4519	ATTGTATTAATTTGTATTAATTTTCCATTTGGAGCTTAACCTGACGCTTGTGTAAGAT	4578
OY	4191	TATGGAAGTGGTAAAGACTG- - - - -AATGTTGAATTAAGTCTAGCTGCTTAAATTAAGTAA	4249
Db	4579	TATGGAAGTGGTAAAGTATTTGAATGTTTGCATTAAGT- - - - -TCTATTAATTAAGTAA	4639
OY	4250	ACTTTTATATCAAAAAAAAAAAAAAAAAAAAA 4279	

```

Db      4636 ACTTTCATGCGAAGAAAAAAGAAAAA 4665
RESULT 12
AR203322
LOCUS      AR203322          4669 bp      DNA      linear      PAT 20-JUN-2002
DEFINITION Sequence 1 from patent US 6365357.
ACCESSION AR203322
VERSION    AR203322.1  GI:21499681
KEYWORDS
SOURCE
ORGANISM   Unknown.
REFERENCE   1 (bases 1 to 4669)
AUTHORS    Mechtner,E. and Fruehauf,J.
TITLE       Methods and reagents for preparing and using immunological agents
            specific for P-glycoprotein
            Patent: US 6365357-A 1 02-APR-2002;
            Location/Qualifiers
                source
                1..4669
                /organism="Unknown"
BASE COUNT 1393 a 894 c 1130 g 1252 t
ORIGIN
Query Match      77.9%; Score 333.2; DB 6; Length 4669;
Best Local Similarity 87.5%; Fred. No. 0;
Matches 3753; Conservative 0; Mismatches 493; Indels 44; Gaps 8;

QY      1  GGAAGCCGAGGTGCGGATGATCCGTAAGAGAGCGGCTAAGGGAGTGCAGAGAAGAACTT 60
Db      409  GGAAGCCGAGGTGCGGATGATCCGTAAGAGAGCGGACCCGAATGAGAGCAAGAAAGAA 468
QY      61  CTGGAAGATGGGCAAAAAAGTAAAAAATGAGAGAGAAAGAAAGCAACTGTCAG 120
Db      469  CTTTTTTAACTGAACATAAAGTGAAGATGAAGAAAGAAAGAAAGCAACTGTCAG 528
QY      121  CACGTTTGCAATGTTCCGCTATTCAAATGCGTTGATAGTTTATATGTTGGTGGGAC 180
Db      529  TGTATTTTCAATGTTTGGCTATTCAAATGCGTTGATAGTTTATATGTTGGTGGGAC 588
QY      181  AATGCTGCGCATTCATCCATGAGCTGCACCTCTCATGATGCTGGTTTGGAAACAT 240
Db      589  TTTGGCTGCGCATTCATCCATGAGGCTGCACCTCTCATGATGCTGGTTTGGGAAAT 648
QY      241  GACAGATAGCTTTGCAATGAGAGAAATTCAGAGAAACAAACTTTTCCAGTTATATTAA 300
Db      649  GACAGATATCTTTCGAATGAGAGAAATTTAGAGATCTGATGT-----CAACATCAC 702
QY      301  TGAAGATATTCAGAACATACACACATTTTCATCAACCATCTGAGAGAGAAATGACAC 360
Db      703  TAATGAGATGATATCAATGATACAGGCTTTCATGATGATCGAGAGAGAAATGACAC 762
QY      361  GTATGCTATTTATTCAGATGAGTGGTGGCTGGCGTGGCTGGCTGCTTATCATCCAGGT 420
Db      763  GTATGCTATTTATTCAGATGAGTGGTGGCTGGCTGGCTGGCTGCTTATCATTCAGGT 822
QY      421  TTCAATTCGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCT 480
Db      823  TTCAATTCGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCT 882
QY      481  TGTATCATGCGACAGAGATTTGGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCTGG 540
Db      883  TGTATCATGCGACAGAGATTTGGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCTGG 600
QY      541  CCGGCTCAAGACGATGCTCCAAATCAATGAAGAAATTTGGGACAAATTTGGAATGTT 1002
Db      943  CCGACTTAAAGATGATGCTCCAAAGATTAATGAAGAAATTTGGGACAAATTTGGAATGTT 1002
QY      601  CTTTCACTCAATAGCAACATTTTACCGGTTTATATGTTGGGTTTAAACAAGTGGTTGAA 660
Db      1003  CTTTCACTCAATAGCAACATTTTACCGGTTTATATGTTGGGTTTAAACAAGTGGTTGAA 1062

```

```

QY      661  GCTAACCTTTGATATTTTGGCCATCAGCCCTGTTCTTGAGCTTTACGCCCATCTGGGC 720
Db      1063  GCTAACCTTTGATATTTTGGCCATCAGCCCTGTTCTTGAGCTTTACGCCCATCTGGGC 1122
QY      721  AAGATATCATCTCATTTTACGATTAAGAACTCTTGGCTATGCAAAAGCTGAGCAGT 780
Db      1123  AAGATATCATCTCATTTTACGATTAAGAACTCTTGGCTATGCAAAAGCTGAGCAGT 1182
QY      781  AGCTGAAGAACTTATGACGAATCAGAACTGATTTGCCCTTGGAGAGCAAAAGAAAGA 840
Db      1183  AGCTGAAGAACTTATGACGAATCAGAACTGATTTGCCCTTGGAGAGCAAAAGAAAGA 1242
QY      841  ACTTGAAGATCAACAAAAATTTAGAAAGCTTAAAGAAATTTGGATTAAGAAAGCTAT 900
Db      1243  ACTTGAAGATCAACAAAAATTTAGAAAGCTTAAAGAAATTTGGATTAAGAAAGCTAT 1302
QY      901  CAGGCGCAACATTTCTATTTGCTGCGCTTCTTATTGATATGATCATCATATGCTGGC 960
Db      1303  TACAGCCAAATTTCTATTTGCTGCGCTTCTTATTGATATGATCATCATATGCTGGC 1362
QY      961  TTTCTGATGAGGACCTCTTGGCTCTCTCAGTGAATATTTATTTGACAAATGCTCAC 1020
Db      1363  CTTCTGATGAGGACCTCTTGGCTCTCTCAGTGAATATTTATTTGACAAATGCTCAC 1422
QY      1021  TGTCTTCTTTCTGATTTAATTTGGGCTTTTATGATTTGAGAGAGATCCCAAGCTTGA 1080
Db      1423  TGTATCTTTCTGATTTAATTTGGGCTTTTATGATTTGAGAGAGATCCCAAGCTTGA 1482
QY      1081  AGCATTTGCAAGAGCAAGAGAGACACTTATGAATCTTCAAGATTAATTTGAATAAC 1140
Db      1483  AGCATTTGCAAGAGCAAGAGAGACACTTATGAATCTTCAAGATTAATTTGAATAAC 1542
QY      1141  AAGCATTGACAGCTATTTGACAGATGAGACATTAACCATATATTTAAGGAAATTTGGA 1200
Db      1543  AAGCATTTGACAGCTATTTGACAGATGAGGACCAACCATATATTTAAGGAAATTTGGA 1602
QY      1201  ATTCAAAATATGTTCACTTATGATTTACCTTCTCGAAAAGAACTTAAAGTCTTAAAGCT 1260
Db      1603  ATTCAAAATATGTTCACTTATGATTTACCTTCTCGAAAAGAACTTAAAGTCTTAAAGCT 1662
QY      1261  CAACCTGAAGGTTTCAGAGTGGGAGACAGAGTGGGCTGGTTGGGAAACAGTGGCGGAA 1320
Db      1663  GAACCTGAAGGTTTCAGAGTGGGAGACAGAGTGGGCTGGTTGGGAAACAGTGGCGGAA 1722
QY      1321  GAGCAGACAGCGTGGCTGATGACAGAGGCTCTATGACCCCAAGATGAGCTGCTGAT 1380
Db      1723  GAGCAGACAGCGTGGCTGATGACAGAGGCTCTATGACCCCAAGAGGATGCTGCTGAT 1782
QY      1381  TGATGACAGAGCATTTAGACATTAATTTAAGCATCTTGGGAAATTTACATGCTGTGT 1440
Db      1783  TGATGACAGAGCATTTAGACATTAATTTAAGCATCTTGGGAAATTTACATGCTGTGT 1842
QY      1441  GAGTGAAGAGCCTGTGTTGTTGCGACACAGATTAAGTGAAGAACTTGGTATGCGCTG 1500
Db      1843  GAGTGAAGAGCCTGTGTTGTTGCGACACAGATTAAGTGAAGAACTTGGTATGCGCTG 1902
QY      1501  AATGTCACATGATGATGATTTGAGAAAGCTGTTAAGAAAGCAATGCTATGATTTAT 1560
Db      1903  AATGTCACATGATGATGATTTGAGAAAGCTGTTAAGAAAGCAATGCTATGATTTAT 1962
QY      1561  CATGAACATCTTAATTAATTTGACACTGTGTTGAGAGAGAGGCGGCAAGTTGAGTGG 1620
Db      1963  CATGAACATCTCTTAATTAATTTGACACTGTGTTGAGAGAGAGGCGGCAAGTTGAGTGG 2022
QY      1621  TGGACAGAAACAGAGATGGCAATTTGCGGCGCTGCTGTTGCAACCCCAAGATTTCTT 1680
Db      2023  TGGACAGAAACAGAGATGGCAATTTGCGGCGCTGCTGTTGCAACCCCAAGATTTCTT 2082
QY      1681  GCTTGATGAGGCAACGTCAGCTTGTGAGACTGAAGTGAACAGAGGTTCAAGTGGGCTT 1740
Db      2083  GCTTGATGAGGCAACGTCAGCTTGTGAGACTGAAGTGAACAGAGGTTCAAGTGGGCTT 2142
QY      1741  GGATTAAGGCCAGAAAAGCGCGGACTACCATTTGTATGCTCATGCTTGTCTACAGTTGG 1800

```

|||||
Db 2143 GGATAGGCGGAAAGGTTGGAGACCACTTGTGATGCTCATCGTTTTGCTACAGTTTG 2202
1801 TATGCCGATGTCATTCGTGTTTGTATGATGAGTCATTTGTGAGAAAGAAATCATGA 1860
2203 TATGCTGACGATCGCTGCTGTTTCGATGATGAGTCATTTGTGAGAAAGAAATCATGA 2262
1861 TGACATCAGAAAGAAAGGAGGATTTACTTCAAACTTGTCAAAATGACAGACAGAGAA 1920
2263 TGACATCAGAAAGAAAGGAGGATTTACTTCAAACTTGTCAAAATGACAGACAGAGAA 2322
1921 TGAATTTGAGTTAGAAATGCGCTGTGATCCAAAAGTGAAGTATGCTTGGAAAT 1980
2232 TGAAGTTGATTTAGAAATGCGCTGTGATCCAAAAGTGAAGTATGCTTGGAAAT 2382
1981 GTCTCCAAAAGATTCAGGGTCCAGTTTAAATAAAGAGATCAATCGCAGAGTATACA 2040
2383 GTCTTCAAAATGATTCAGATCCAGTCTAAATGAAGAAAGATCACTGATGAGAGTGTCCG 2442
2041 TGCACCAAGGCCAAGACAGAAAGCTGTGTACAAAAGAGACTTGAATGAGATGACC 2100
2443 TGGATCACAGCCCAACACAGAAAGCTGTGTACAAAAGAGACTTGAATGAGATGACC 2502
2101 TCCAGTTTCTCTCGAGAGATTCAGAGCTGAAGCTCAACTGAATGAGCTTATTTTGTGCT 2160
2503 TCCAGTTTCTCTCGAGAGATTCAGAGCTGAAGCTCAACTGAATGAGCTTATTTTGTGCT 2562
2161 TGGATATTTTGTCTATTTAAACGAGAGCTGCACACAGATTTCAATATATTTTTC 2220
2563 TGGTGTATTTTGTGCTATTTAAATGAGAGCTGCACACAGATTTCAATATATTTTTC 2622
2221 AAGGATTTAGGAGTCTTTACCCGATGAGATGCTGAAGCTGAAACGACAGATATGTA 2280
2623 AAGGATTTAGGAGTCTTTACCAAGATTTGATTCGAAACGACAGATATGTA 2682
2281 CAGTTTCTGTATTTGTTCTAGTCTGTGATATTTCTTTATTAATGATTTTCTCTCA 2340
2683 CTGTGTTTCTGTATTTGTTCTAGTCTGTGATATTTCTTTATTAATGATTTTCTCTCA 2742
2341 GGGCTTCACATTTGGCAAGGCTGGGAGAGTCTCACTAGCGGCTTCGATATCGTGTTC 2400
2743 GGGCTTCACATTTGGCAAGGCTGGGAGAGTCTCACTAGCGGCTTCGATATCGTGTTC 2802
2401 CAGATTCATGCTGAGACAGAGATGTCAGTGTGATGATGACCTTAAACGACACTGAGAG 2460
2803 CCGATTCATGCTGAGACAGAGATGTCAGTGTGATGATGACCTTAAACGACACTGAGAG 2862
2461 ATTGACACACAGGCTTGGCAATGATGCGGCTCAAGTTTAAAGGGGCTATAGGTTCCAGGCT 2520
2863 ATTGACACACAGGCTTGGCAATGATGCGGCTCAAGTTTAAAGGGGCTATAGGTTCCAGGCT 2922
2521 TGGTGTATTTACCCAGAAATATGACAAATCTTGGGACAGGCACTTATATCTTAAATCTA 2580
2923 TGGTGTATTTACCCAGAAATATGACAAATCTTGGGACAGGCACTTATATCTTAAATCTA 2982
2581 TGGTGTGCAATTTAACTTTACTTGTAGCAATTTGTACCAATTCGATATGACAGAGAT 2640
2983 TGGTGTGCAATTTAACTTTACTTGTAGCAATTTGTACCAATTCGATATGACAGAGAT 3042
2641 TGTGAAATGAAATGTTGTCTGAGACAGACTGAAAGATTAAGAAAGGCTATGAAAGGAGC 2700
3043 TGTGAAATGAAATGTTGTCTGAGACAGACTGAAAGATTAAGAAAGGCTATGAAAGGAGC 3102
2701 TGGGAGATTTCTACAGAAACCATCGAAACCTTCGAACTGTGTTTCTTTCAGCTGAGGA 2760
3103 TGGGAGATTTCTACAGAAACCATCGAAACCTTCGAACTGTGTTTCTTTCAGCTGAGGA 3162
2761 GCGAAGATTTGAATACATGATGACAGAGTGTTCAGAGTACATACAGAAACCTTTCAG 2820
3163 GCGAAGATTTGAATACATGATGATGACAGAGTGTTCAGAGTACATACAGAAACCTTTCAG 3222
2821 GAAAGCACACATCTTGGGGCTCATTTTCTATCACCCAGCAATGATATATTTTCTCA 2880
|||||

Db 3223 GAAAGCACACATCTTTGGATTTACATTTTCTTCCCTTCACCCAGGCAATGATATTTTCTCA 3282
2881 TGGTGTGCTTTCCGGTTTGTGATGCTTACTGTTGGCAAAATGATTCATGAACTTCAGGA 2940
3283 TGGTGTGATGTTCCGGTTTGTGAGCTTACTGTTGGCAAAATGATTCATGAACTTCAGGA 3342
2941 TGTCTTTTGTGATTTCTGAGCTATGCTTTGGTGTGCTATGAGAGTGTGAGGAGTCACTTC 3000
3343 TGTCTTTTGTGATTTCTGAGCTATGCTTTGGTGTGCTATGAGAGTGTGAGGAGTCACTTC 3402
3001 ATTGCTCTGATATGCAAAAGCCAAAGTATACAGAGCCCACTGATCATGATCATTTGA 3060
3403 ATTGCTCTGATATGCAAAAGCCAAAGTATACAGAGCCCACTGATCATGATCATTTGA 3462
3061 AAAAGCCCTCTGATTTGACAGCTTACAGCCCTTCAGAGGCTTCAGAAATTCAGTTGAGAG 3120
3463 AAAAGCCCTTGTGATTTGACAGCTTACAGCCCTTCAGAGGCTTCAGAAATTCAGTTGAGAG 3522
3121 AAATGTGATTTATGATGCTGTGTTCAACTATCCACTGACCAAGCATCCCGTGTCT 3180
3523 AAATGTGATTTGATGATGCTGTGTTCAACTATCCACTGACCAAGCATCCCGTGTCT 3582
3181 CCAGGCTGAGCTCCAGATGAAAGAGGCGCAGAGCTGAGGCTTCGATGAGCAATGG 3240
3583 TCAGGAGTGTAGGCTGTGAGGAGTGAAGAGGCGCAGAGCTGAGGCTTCGATGAGCAATGG 3642
3241 CTGTGGAAAGACACAGTTGTTCACTCTAGAGCGCTTCATGAGCCCTTGGCTGTGTTTC 3300
3643 CTGTGGAAAGACACAGTTGTTCACTCTAGAGCGGTTCTAGAGCCCTTGGCTGTGTTTC 3702
3301 AGTGTATTTGATGAGCAAGATTAAGCAACCTGATGCTGCTGCTCCGAGACACCT 3360
3703 AGTGTATTTGATGAGCAAGATTAAGCAACCTGATGCTGCTGCTCCGAGACACCT 3762
3361 GGGCATGCTGTGTCAGAGAGCCATCTGTTGATGCTGACATTTGCGAGAAATTCGCTA 3420
3763 GGGCATGCTGTGTCAGAGAGCCATCTGTTGATGCTGACATTTGCGAGAAATTCGCTA 3822
3421 TGGAGACACAGCGGCTGTATCATGATGAGATTAATGAGGAGCAGCAGCAAGAGGCCAA 3480
3823 TGGAGACACAGCGGCTGTATCATGATGAGATTAATGAGGAGCAGCAGCAAGAGGCCAA 3882
3481 CATACACACATTTGATGAGACACCTCCCTGAGAAATATACACACAGATTAAGGAGCAAAAG 3540
3883 CATACACACATTTGATGAGACACCTCCCTGAGAAATATACACACAGATTAAGGAGCAAAAG 3942
3541 AACCCAGCTCTGTGTTGGCCAGAAACAGGCAATGCTGATGCTGCTGTTTGAACA 3600
3943 AACCCAGCTCTGTGTTGGCCAGAAACAGGCAATGCTGATGCTGCTGTTTGAACA 4002
3601 GCTCATATTTTGTGTTTGGATGAGAGCTACATGAGCTCTGATTCAGAAATGAGAAAGT 3660
4003 GCTCATATTTTGTGTTTGGATGAGAGCTACATGAGCTCTGATTCAGAAATGAGAAAGT 4062
3661 TGTCCAGAGCCCTGGACAAAGCCAGAGAGGCGCACCTGATTTGATGCTCCACCG 3720
4063 TGTCCAGAGCCCTGGACAAAGCCAGAGAGGCGCACCTGATTTGATGCTCCACCG 4122
3721 CTGTGTCACATTCAGAGATGAGATTAATGATGTTGTTTCAAGATGCAAGTCAAGGA 3780
4123 CTGTGTCACATTCAGAGATGAGATTAATGATGTTGTTTCAAGATGCAAGTCAAGGA 4182
3781 GCATGGACACATCAACAGCTGCTGGCCAGAAAGGATATTTTCCATGCTGAGTGT 3840
4183 GCATGGACACATCAACAGCTGCTGGCCAGAAAGGATATTTTCCATGCTGAGTGT 4242
3841 CCAGGCTGAGAAAGCCCTAGTAACTGTGCTATGAGAGTGTAAATATTTTAAAT 3900
4243 CCAGGCTGAGAAAGCCCTAGTAACTGTGCTATGAGAGTGTAAATATTTTAAAT 4302
3901 ATTGCTTAAACATGAGCATTTTAAATCAAAATGAGGAGGAGCACTTTCGAGAAATCT 3960
4303 ATT - TGTGATATGATGATTTTAAAGTTAA - AGCAACACATTCAGAGATTT 4355
|||||

RESULT	13
AX391099	
LOCUS	AX391099
DEFINITION	AX391099
ACCESSION	Sequence 1 from Patent WO0210205.
VERSION	AX391099
KEYWORDS	AX391099.1 GI:19584228
SOURCE	.
ORGANISM	human. Homo sapiens
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
AUTHORS	Ruth,A. and Roulnson,I.
TITLE	Mutations of the mdrl p-glycoprotein that improve its ability to confert resistance to chemotherapeutic drugs
JOURNAL	Patent: WO 0210205-A 1 07-FEB-2002.
FEATURES	THE BOARD OF TRUSTEES OF THE UNIVERSITY OF ILLINOIS (US) Location/Qualifiers

CDS

/*organism="Homo sapiens"
/db_xref="taxon:9606"
425..44267
/note="unamed protein product"
/codon_start=1
/protein_id="CAD28599.1"
/db_xref="GI:19584229"
/translation="MDLEGDRLGKAKKFKLNKSEDKKKKPPVSYFSEKRYSN
MDLKLVMYVGLTAIIHGAGLPIMLVNGEMTDIPANAGNIDIMKSTNIRSDINDTG
FPAHNLDEMDERARYYSGISGAGVIAATIVQSECLAGNIDIKRQFHAIMRDEI
GMFDVWDEGELNTRLRLDYVSKNIENSGIKTGMFQSAETFEFGIVTFRGKMLTLEI
LASPLIGLSAAVMKAIISSEFDEKLVAAGAQADEVIAAARTVIAFGSGKTELLEI
YRNLEAKRIGIKKAITANISIGAFELIYASVALFWYGLTVLVSSEKISGOIVATY
FSEVLLGASVQASPSIEAFANARGAAYELFTIKDKPSIDISYSGKSPKNTINGNL
EENFNHYSYPSREVKILKNGKQSOQVYALGNSGCGKSTYVOMLRDLPDEGM
VSDGODDIRTIWVRLRELIGVYQSEPLTATENIRGENTVMDLRLPDEGM
AVDEIKLPHKFDTLVBERGAOLISGQKQALATARALVRNPKILLDEATSAIDTSE
AAVOVALADGKRGRTTVIAHRLSTVRNADVIAGFDGQVVENGNHDELMKRGTYFK
LYTMQDAGSEVLEMAADESKSEIDALEMSNDRSRLKRSNRSRVSGQAQDRKL
SKRLADESIPEVFWRIKMLNIIISPELFGQFFEGAGELITRLERLYVERSMR
QVSWEPDEPKNTIAGTLNRLANDAAQYKAGISRLVATQNTANLGTGIIISFYQEO
LTLTLALVPIIAGAVGEMKMLSGQALDKRELKAGATAPETAIENPETYVSLQEO
KREHNYAQSLOVPRNSLRKAHIGIIFSTFQTMQMAEVSQAQCPGCAIVAKKMSFE
DVLTAFSVAEGMAAGVQSEFADYKAKTSAHAHITIKETPLIDISTIDGLMPNT
LEGANTFEEVYENTPTBDDIPIVGLQSLTPKKQOTLALVGGSGGCKSTVQLLEFED
PLAGVILIDGKEIKRLVOMIRAHGLTVSEDPISIAENIAGDNRVYVQSEIY
RAAKNANHATIESLPKYSTKVDGKQTLISGQKQALAIATVLRQPHIILLDEATS

BASE COUNT	1372 a	892 c	1130 g	1252 t
ORIGIN	AADTSEKVOEALDKAREGRTCIIVIAHRLSTIQNADLIYFQNGRVKEHGTHOOLLA QNGIFSMVSQAQGRQ"			

Query Match	77.8%;	Score 3327.6;	DB 6;	Length 4646;
Best Local Similarity	87.6%;	Pred. No. 0;		
Matches 3742;	Conservative	0;	Mismatches 484.	Indels 44.

QY	1	GGAGCGCAGAGTGGAGTGGATCCGTGAAGAGAGCGCTTAAGAGGAGTGCAGAGAAACCTT	60
Db	409	GGAGGCGCAGAGTGGAGTGGATCTTGAAGGGGACCGCAATGAGAGAGCAAAAGAGAA	46
QY	61	CTGGAAAAATGGGCAAAAAAATATAAAAAATGAGACAAAGAAAAGAAACCAACTGTGAG	12
Db	469	CTTTTCTTTTAACTGCAACATTAATAAGTAAAAAGATTAAGAAAGAAAAGAAACCAACTGTGAG	52
QY	121	CACGTTTGCAATGTTCGTCTTCAATTAATTTGGCTTGATPAGTGGTATATGTGTGGGAC	18
Db	529	TGTATTTTCAATGTTCGTCTTCAATTTGGCTTGACAAAGTTGTATATGTGTGGGAC	58
QY	161	AATGGCTGCCATCATTCATGAGAGCTGCATCCCTCTCATGATGCTGGTTTGGAAACAT	24
Db	569	TTTGGCTGCCATCATTCATGAGAGCTGCATCCCTCTCATGATGCTGGTTTGGAAACAT	64
QY	241	GACAGATAGCTTGCCAAATGACAGGAATTCAGAAAACAAACTTTCCAGTTATATTTAA	30
Db	649	GACAGATATCTTGCCAAATGACAGGAATTTAAGAAAGATGATGTCA-----ACATTCAC	70
QY	301	TGAAGATTATACGAAATATACAAATTTACAAATTCACCAATCTGGAGAGGAAATGACCAC	36
Db	703	TAAATGAAGATATCATGATATCAGAGGTTCTTCAATGATGAGAGGAAAGATGACCAG	76
QY	361	GTAAGCTATTTATACATGAGAGAGTGGCTGGCGCTGCTGAGGCTGCTTATCCACAGT	42
Db	763	GTAAGCTATTTATACATGAGAGATGGTCTGGCTGCTGAGGCTGCTTATCCATCCAGT	82
QY	421	TTCAATTCGTGCTGGCGAGCGAGAAAGACAGTACTCAAAATTAGAAAACAAATTTTCA	48
Db	823	TTCAATTTTGGTCTGGCGAGCGAGAAAGACAAATATACAAAAATTAGAAAACAGTTTTC	88
QY	481	TGCTATATGCGACAGAGATTTGGCTGTTGACGTGATACGTGGGGAGCTTAACAC	54
Db	883	TGCTATATGCGACAGAGATTTGGCTGTTGACGTGATACGTGGGGAGCTTAACAC	94
QY	541	CGGCGTCACAGACAGATGTCTCCAAATCATTAAGAGGAATTTGGCGACAAATTTGGAATGT	60
Db	943	CGGCGTCACAGATGTCTCTAAATTAAGAGGAATTTGGCGACAAATTTGGAATGT	100
QY	601	CTTTCACATATAGCAACATTTTCAACGGTTTATATGTGGGGTTTACAGTGGTTGAA	66
Db	1003	CTTTCAGTATAGCAACATTTTCAACGGTTTATATGTGGGGTTTACAGTGGTTGAA	106
QY	661	GCTAAACCTTGATTTTGGCCATCAGCCCTGTCTTGAGCTTTCAGCCGCCATCGGGC	72
Db	1063	GCTAAACCTTGATTTTGGCCATCAGCCCTGTCTTGAGCTTTCAGCCGCCATCGGGC	112
QY	721	AAAGATATCTATCTCATTTATATGATTAAGAACTTTGGCCATGACAAAAGCTGAGCAGT	78
Db	1123	AAAGATATCTATCTCATTTATGATTAAGAACTTTGAGGTATGCAAAAGCTGAGCAGT	118
QY	781	AGCTGAAGAGCTCTTAGCAGCAATCAGACTGATTTGGCTTTGAGAGCAAAAGAAAGA	84
Db	1183	AGCTGAAGAGCTCTTAGCAGCAATTAGAAGCTGATTTGAGAGCAAAAGAAAGA	124
QY	841	ACTTGAAGAGTACACAAAAATTTAGAGAGCTTAAAGGAATTTGGGATTAAGAAAGCTAT	90
Db	1243	ACTTGAAGAGTACACAAAAATTTAGAGAGCTTAAAGGAATTTGGGATTAAGAAAGCTAT	130
QY	901	CACGCGCAACATTTCTATTTGGCGGCTTCTTATGATCTATGATCATATGCTGTGC	96
Db	1303	TACAGCCAAATTTCTATATGAGTGGCTGCTTCTGCTGATCATATGATCTTATGCTGTGC	136

OY	961	TTTCTGGATATGGAACCTTCCTTGGCTCTCTCCATGTAATATTTCTATATGAGCAACTACTCAC	1020
Db	1363	CTTCTGGATATGGAACCACTTGGCTCTCTCCATGAGGAATATTTCTATATGAGCAACTACTCAC	1422
OY	1021	TGTCTCTTTCTCTGTATTAATATTTGGGGCTTTTATGATTTGAGCAGAGCATCCCAAGCATTTGA	1080
Db	1423	TGTATTTCTTTCTGTATTAATATTTGGGGCTTTTATGATTTGAGCAGAGCATCTCCAAAGCATTTGA	1482
OY	1081	AGCATTTTGCAAAACCCAGAGAGAGCAGCTTATGAATCTTCAAGATTAATTGACAAATAAAC	1140
Db	1483	ACCATTTTGCAAAATGCAGAGAGAGCAGCTTATGAATCTTCAAGATTAATTGATTAATAAGCC	1542
OY	1141	AAGCATTTGACACCTTTTGGAGAGAGAGCATAAACCAAGATATATTTAAGGAAAAATTTTGA	1200
Db	1543	AAGTATTTGACACCTTATTCGAGAGAGAGCATAAACCAAGATATATTTAAGGAAAAATTTTGA	1602
OY	1201	ATTTCAAAAATTTTCACTTCAGTTACGTTTCTCGAAAAGAAAGTTAAGATCTTAAAGGGGCT	1260
Db	1603	ATTTCAGAAAATTTTCACTTCAGTTACGTTTCTCGAAAAGAAATTTAAGATCTTAAAGGGGCT	1662
OY	1261	CAACCTGAAAGTTTCAGAGTGGGGCAGACAGTGGCGCTGGTTGGGAAACAGTGGCTCGGGAA	1320
Db	1663	GAACCTGAAAGTGGCAGAGTGGGGCAGACAGTGGCGCTGGTTGGGAAACAGTGGCTGGGGAA	1722
OY	1321	GAGCAGACCGCTGACACCTGATGCAAGGCTATATGACCCACAGATTTGGCATAGTCTGTAT	1380
Db	1723	GAGCAGACACAGTCCACACTGATGCAAGGCTATATGACCCACAGAGGGAGTGGTACGTGT	1782
OY	1381	TGATGTGACAGAGCATTTAAGGACCACTTAATTAAGGCACTCTTCGGGAAATTTACTGTGTGT	1440
Db	1783	TGATGTGACAGAGATTTAAGGACCACTTAATTAAGGTTTCTACGGGAAATCATGTGTGTGT	1842
OY	1441	GAGTCAGAGACCTGTGTGTTTCCACACACAGATAGCTGAAAACATTCGTATAGGCCCGGA	1500
Db	1843	GAGTCAGAGACCTGTATTTGTTTCCACACAGATAGCTGAAAACATTCGTATAGGCCCGTGA	1902
OY	1501	AAATGTCCACATGGATGATGATGAGAAAGGTGTTAAGGAAGCCATAGCTATGATTTAT	1560
Db	1903	AAATGTCCACATGGATGATGATGAGAAAGGTGTCMAAGGAAGCCATAGCTATGATTTAT	1962
OY	1561	CATGAACCTACCTATTAATTAATTTGACACTCTGTGTTGAGAGAGAGGGCCACGCTGATG	1620
Db	1963	CATGAACCTGCTCATTAATTAATTTGACACCTGTGTTGAGAGAGAGGGCCACAGTTGATG	2022
OY	1621	TGACAGAGAAACAGAAATGCGCATTTGTCGGGCCCTGTTGCGAACCCCAAGATTTCTCT	1680
Db	2023	TGGCGAGAGAGAGAGGATCGCATTTGCAACGTGCCCTGTGCGAACCCCAAGATTTCTCT	2082
OY	1681	GCTGGATAGGCAACGTCAGCTCTGCACTGCACTGAAAGTGAAGCAGTGTATAGGTGGCCCT	1740
Db	2083	GCTGGATAGGGCACGTCAGCTCTTGGCACAGAAAGGAAAGCAGTGTATAGGTGGGCTCT	2142
OY	1741	GGATTAAGCCAGAAAAGGCCCGGACATACCATTGTATAGCTCATAGTTTGTCTACAGTTTG	1800
Db	2143	GGATTAAGGCCAGAAAAGGTGCGACACCATTTGTATAGCTCATAGTTTGTCTACAGTTTG	2202
OY	1801	TATATCCGATGTCTATTGCTGTGTTTATGATGATGAGTCTATTGTGAGAGAAAGAAATCTATGA	1860
Db	2203	TATATGCTGACGTCTATCGCTGTGTTTATGATGATGAGTCTATTGTGAGAGAAAGAAATCTATGA	2262
OY	1861	TGTAATCTATGAAGAGAGAGGCAATTTACTTCAACTTTCACAATTCGACAAATGAGGAAAA	1920
Db	2263	TGTAATCTATGAAGAGAGAGAGGCAATTTACTTCAACTTTCACAATTCGACAAATGAGGAAAA	2322
OY	1921	TGAAATTTGATTAAGAAAATGCACTGCTGTAATCCAAAAGTGAAGTGCCTTTGGAAT	1980
Db	2323	TGAAATTTGATTTGAAAATGCACTGCTGTAATCCAAAAGTGAATTTGATGCTTTGGAAT	2382
OY	1981	GTTCTCCAAAAGATTTCAAGGTCACGTTTAATTAAGAAAGATTCACATCGCAGAGATATCA	2040
Db	2383	GTTCTCAAAATGATTTCAAGATTCACGTTTAATTAAGAAAGATTCACATCGCTAGAGATGTCG	2442
OY	2041	TGCACCAAGCAAGCCAGACAGAAAGTGTGTCAAAAGAGCACTTGAATGGAATGTACC	2100

Db	2443	TGATCAACAAGCCCAAGACACAAGAGCTTACGACCAAGAGGCTCGATTAAGAAAGTAAATACC	2502
Qy	2101	TCGAGTTTCCTCTGAGAGATTCGAAAGCTCAACTGCACTGAAATGGCTTAATTTTGtGT	2160
Db	2503	TCGAGTTTCCTTTGAGAGATTAAGAAAGCTAAATTTAACGAATGGCCCTTAATTTTGtGT	2562
Qy	2161	TGATATATTTTGTCTATTAATTAACGAGAGCTCGCAACCAAGCATTTTCAATTAATTTTC	2220
Db	2563	TGCTGTATTTTGTGACATTAATAAGAGGCTCGCAACGCAATTTGCAATTAATTTTC	2622
Qy	2221	AAGATTAATGGAGATCTTACCCGAGATGAGATCTCGAAACCAACGACAGATAGTAA	2280
Db	2623	AAAGATTAATGGAGCTTTTACAGAAATTTGATGATCTCGAAACCAACGACAGAAATAGTAA	2682
Qy	2281	CATGTTTTCGTATTTCTATTTCTAGTCTTGGAATTAATTTCTTTATTAACATTTTCTCCA	2340
Db	2683	CTTGTTTTCATCTATTGTTTCTGTAGCCCTTGGAATTAATTTCTTTATTAACATTTTCTCCA	2742
Qy	2341	GGGCTTCAACATTTGGCAAAAGCTGGGAGATCCTCACTAAGGGCTTCGATACATGGTTT	2400
Db	2743	GGGTTTCACTTTGGCAAAAGCTGAGAAATCCTCAACAAAGGGCTCGAATAGATGGTTT	2802
Qy	2401	CAGATCATCTGTGACAGAGAGATGCTAGCTGTTTATGATCCCTAAAAACCAACATGAGAC	2460
Db	2803	CGATCCATCTGTGACAGAGAGATGTGATGGTTGATATCCCTTAACCAACCACTGAGAC	2862
Qy	2461	ATTGACCAACGAGGTTTCCCAATGATCGGCTCAAGTTAAAGGGCTTAATGTTCCAGGCT	2520
Db	2863	ATTGATCTACGAGGCTCCCAATGATCTGCTCAAGTTAAAGGGCTTAATGTTCCAGGCT	2922
Qy	2521	TGCTGTCAATACCAAGAAATATACAAATCTTGGGACAGGCAATTTATATCTTAATCTTA	2580
Db	2923	TGCTGTAAATATCCCAAGAAATATACAAATCTTGGGACAGGCAATTTATATCTTCACTTA	2982
Qy	2581	TGTTGGCAATTAACACTTTTACTCTTAGCAATTTGATCCCATTCATTCGAATGACAGAGT	2640
Db	2983	TGTTTGGCACTAACACTCTTACTCTTAGCAATTTGATCCCATTCGAATGACAGAGAGT	3042
Qy	2641	TGTTGAATGAAGAAATGTTGTCTGACACAGCACTGAAAGATTAAGAAAGGATGAGGAG	2700
Db	3043	TGTTGAATGAAGAAATGTTGTCTGACACAGCACTGAAAGATTAAGAAAGACTAGAAAGTGC	3102
Qy	2701	TGGGAGATTTGCTACAGAAAGCATTCGAAACTTCCGAACTGTGTTTCTTTGACTGCGGA	2760
Db	3103	TGGGAGATTTGCTACAGAAAGCATTCGAAACTTCGAAACTGTGTTTCTTTGACTCGAGA	3162
Qy	2761	GCAGAGTTTGAATACATGATATGACACAGAGTTTTCAGATGCCATACAAAACTGTTAG	2820
Db	3163	GCAGAGTTTGAACATATGATATGCTCAGAGATTTTCAGATGCCATACAAAACTGTTAG	3222
Qy	2821	GAAAGCAACATCTTCGGGGTCTAATTTTCTATCACCAGGCAATGATGATATTTTCCCTA	2880
Db	3223	GAAAGCAACATCTTTTGGAAATTAACATTTTCTTCCACAGGCAATGATGATATTTTCCCTA	3282
Qy	2881	TGCTGAGCTGTTTCCGGTTTGTGCTCTACTTGTGGCAATGAGTTTCAAGCATTTTCAGAGA	2940
Db	3283	TGCTGAGATGTTTCCGGTTTGTGAGCTCTACTTGTGGCAATTAATCTATAGCTTTGAGAGA	3342
Qy	2941	TGTTCTTTTGGTATTTCTGACGAAATGTTTGGTGCCATGGCATGGAGTGGGCAAGTCAGTTC	3000
Db	3343	TGTTCTTTTGGTATTTTCAAGCTGTGTTCTTTTGGTGCCATGGCCGTGGGCAAGTCAGTTC	3402
Qy	3001	ATTGTCCTGACTATGCCAAAGCCAAAGATATCAGCAAGCCCACTGATCATATGATCAATTGA	3060
Db	3403	ATTGTCCTGACTATGCCAAAGCCAAATATCAGCAAGCCCACTGATCATATGATCAATTGA	3462
Qy	3061	AAAAAGCCCTCTGATTTGACAGCTACAGCCCTTCAGGGCTCTAAGGCAAAATACGTTGGAGG	3120
Db	3463	AAAAAGCCCTCTGATTTGACAGCTACAGCCGGAAGGCTTAATGCCCAAGCAATTTGGAAGG	3522
Qy	3121	AAATGTGACATTTAATAGATCTGTTCACATATCCACTGTGACCAACATCCCGGTGCT	3180

Db 3523 AAATGTCACATTGGTGAAGTGTGTAATTCACATTCGCCACCCGACGACATCCCAAGTCT 3582
Qy 3181 CCAGGGGCTGAGGCTCGAGGTGTAAGAGGGCCAGACGCTGGCCCTCGTAGTAGTACAGTGG 3240
Db 3583 TCAGGAGTGAACCTGGAGGTGAAGAGGGCCAGACGCTGGCTGTGGTGGGACAGTGG 3642
Qy 3241 CTGTGGGAAGACAGTGTTCAGCTCTGAGGCTGAGGCTGTCAATGACCCCTTGGCTGGTTC 3300
Db 3643 CTGTGGGAAGACAGTGTTCAGCTCTGAGGCTGAGGCTGTCAATGACCCCTTGGCTGGTTC 3300
Qy 3301 AGTGTATGTATGAGCAAGAGATTAAGACACCTGTAATGTCCAGTGGCTCCGACACAGCT 3360
Db 3703 AGTGTGTCTGTATGAGCAAGAGATTAAGACACCTGTAATGTCCAGTGGCTCCGACACAGCT 3762
Qy 3361 GGGGATCTGTCTGAGAGAGCCATCTGTTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGT 3420
Db 3763 GGGGATCTGTCTGAGAGAGCCATCTGTTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGT 3822
Qy 3421 TGGAGACACAGCGGGGCTGTATTCATGTAAGAGATTAAGAGAGAGAGAGAGAGAGAGAG 3480
Db 3823 TGGAGACACAGCGGGGCTGTATTCATGTAAGAGATTAAGAGAGAGAGAGAGAGAGAGAG 3882
Qy 3481 CATACACACTGATCGAGACACTCCCTGTAAGATTAAGAGAGAGAGAGAGAGAGAGAGAG 3540
Db 3883 CATACACACTGATCGAGACACTCCCTGTAAGATTAAGAGAGAGAGAGAGAGAGAGAGAG 3942
Qy 3541 AACCCAGCTCTGTGTGGCCAGAAAGCGCATGTCATAGTGTGAGTGTGAGTGTGAGTGTGAGT 3600
Db 3943 AACCCAGCTCTGTGTGGCCAGAAAGCGCATGTCATAGTGTGAGTGTGAGTGTGAGTGTGAGT 4002
Qy 3601 GCCTCATATTTGCTTTGATGTAAGTCAATCAGCTCTGATACAGAAAGTGAAGAGT 3660
Db 4003 GCCTCATATTTGCTTTGATGTAAGTCAATCAGCTCTGATACAGAAAGTGAAGAGT 4062
Qy 3661 TGTCCAAAGAGCCCTGGAGCAAAAGCCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 3720
Db 4063 TGTCCAAAGAGCCCTGGAGCAAAAGCCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 4122
Qy 3721 CTGTGTCACATTCAGAAATGAGATTAATGATGATGATGATGATGATGATGATGATGATGAT 3780
Db 4123 CTGTGTCACATTCAGAAATGAGATTAATGATGATGATGATGATGATGATGATGATGATGAT 4182
Qy 3781 GGATGGCACATCAACAGCTGTGGCCAGAAAGAGATCTATTTTCCAGTGTGAGTGTGAGTGT 3840
Db 4183 GGATGGCACATCAACAGCTGTGGCCAGAAAGAGATCTATTTTCCAGTGTGAGTGTGAGTGT 4242
Qy 3841 CCAGGCTGAGCAAAAGCGCTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGT 3900
Db 4243 CCAGGCTGAGCAAAAGCGCGCTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGT 4302
Qy 3901 ATTGTGTAAACATGAGATTAATCAAGTAAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAG 3960
Db 4303 ATT--TGTTTGAATATGACATTTATTCAGTTAA--ACCAACACTTTCAGAAAT 4355
Qy 3961 ATTGTGAACATGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAG 4020
Db 4356 ATGAAGAGATGCTGTGTAACATTTCTGAGTCAAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAG 4403
Qy 4021 AGTCTTGAATTTTAAATTAAGAACCA--AAGAAATATATCTGAGTGTGAGTGTGAGTGTGAG 4078
Db 4404 ----AGAGACTTCGTAAATTAAGAACCAAGAGTGAAGATCATCATCAAGTGTGAGTGTGAG 4458
Qy 4079 ACTGTGTAAATGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAG 4138
Db 4459 CATAGTTTAACTGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 4518
Qy 4139 TGTATATTTTGTATATATTT-----ATTGTACCTAGCTGTGAGTGTGAGTGTGAGTGTGAG 4190
Db 4519 ATGTGTAAATTTTGTATATTTTCCCATTTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAG 4578
Qy 4191 TATAGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAG 4249
Db 4579 TATAGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAG 4635

Qy 4250 ACTTTTATAT 4259
Db 4636 ACTTTCATGT 4645

RESULT 14
AX336420
LOCUS AX336420
DEFINITION Sequence 6929 from Patent WO0194629. DNA 4646 bp 1 linear PAT 09-JAN-2002
ACCESSION AX336420
VERSION AX336420.1 GI:18127139
KEYWORDS
SOURCE human.
ORGANISM Homo sapiens
REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
AUTHORS Young, P. E., Augustus, M., Carter, K. C., Edner, R., Endress, G., Horigan, S., Soppet, D. R. and Weaver, Z.
TITLE Cancer gene determination and therapeutic screening using signature gene sets
JOURNAL Patent: WO 0194629-A 6929 13-DEC-2001;
FEATURES
source location/Qualifiers
1..4646
BASE COUNT 1371 a 892 c 1129 g 1254 t
ORIGIN

Query Match 77.7%; Score 3324.4; DB 6; Length 4646.
Best Local Similarity 87.6%; Pred. No. 0;
Matches 3740; Conservative 0; Mismatches 486; Indels 44; Gaps 8;

Qy 1 GGAGCGCGAGTGGGATGATCCCTGAAAGAGCGCGTGAAGGGAGTCCAGAGAACTT 60
Db 409 GGAGCGCGAGTGGGATGATCCCTGAAAGAGCGCGTGAAGGGAGTCCAGAGAACTT 60
Qy 61 CTGGAATAATGGGCAAAAAAATGTAATAAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 120
Db 469 CTGGAATAATGGGCAAAAAAATGTAATAAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 120
Qy 121 CACGTTGCAATGTTGCTGCTATTCATTAATGCTTGAATGATGATGATGATGATGATGATGATGAT 180
Db 529 TGTATTTTCAATGTTGCTGCTATTCATTAATGCTTGAATGATGATGATGATGATGATGATGATGAT 588
Qy 181 AATGGCTGCATTCATCCATGAGCTGACATCCCTCATGATGATGATGATGATGATGATGATGATGAT 240
Db 589 TTTGGCTGCATTCATCCATGAGCTGACATCCCTCATGATGATGATGATGATGATGATGATGATGAT 648
Qy 241 GACGATAGCTTTGCAATGAGAGATTTCAAGAAACAAATTTTCCAGTTATATTA 300
Db 649 GACGATAGCTTTGCAATGAGAGATTTCAAGAAACAAATTTTCCAGTTATATTA 300
Qy 301 TGAAGTATTAAGACAAATACACATTTTCATCAACCAATGAGAGAGAGAGAGAGAGAGAGAGAG 360
Db 703 TGAAGTATTAAGACAAATACACATTTTCATCAACCAATGAGAGAGAGAGAGAGAGAGAGAGAG 762
Qy 361 GTATGCTATTAATTAAGAGTGGATCGCTGCGCTGCTGATGATGATGATGATGATGATGATGATGAT 420
Db 763 GTATGCTATTAATTAAGAGTGGATCGCTGCGCTGCTGATGATGATGATGATGATGATGATGATGAT 822
Qy 421 TTGATCTGAGTGGCTGCGAG 480
Db 823 TTGATCTGAGTGGCTGCGAG 882
Qy 481 TGTATATGAGAGAGAGATGCTGTTGAGTGTGATGATGATGATGATGATGATGATGATGATGATGAT 540
Db 883 TGTATATGAGAGAGAGATGCTGTTGAGTGTGATGATGATGATGATGATGATGATGATGATGATGAT 942
Qy 541 CCGGCTCAGAGAGATGCTCCAAATCAATGAAGAGATTTGGCAAAATGAGATGTT 600
Db 942 CCGGCTCAGAGAGATGCTCCAAATCAATGAAGAGATTTGGCAAAATGAGATGTT 600

Db	943	CCGACTTACAGATGATGTCCTCTAAGATTAAATGACGTTATTGGTGCAAAATTTGGAATGTT	1002
Qy	601	CTTTCACATCAATACACACATTTTTCACCGGGTTTTTAATAGTGGGGTTACACGTGGTTGGAA	660
Db	1003	CTTTCAGTCATATGSCAACATTTTTCACCTGGGTTAATAGATTTTACACGCTGGTTGGAA	1062
Qy	661	GCTAACCCCTTGATATTTTGGCCATCAGCCCTGTCTTGGACCTTTCAGCCGCACATCGGGC	720
Db	1063	GCTAACCCCTTGATATTTTGGCCACAGCTCTGTCTTGGACCTTTCAGCTGCTGTGTGGCC	1122
Qy	721	AAAGTACTATCTCATTTACTGATTAAGAAACTCTTGGCCATAGCAAAAGCTGGACACT	780
Db	1123	AAAGTACTATCTCATTTACTGATTAAGAAACTCTTGGCCATAGCAAAAGCTGGACACT	1182
Qy	781	AGCTGAAGAAGTCTTATGACAGAAATCAGACCTGTGATTCCTTTGGAGAGCAAAAGAAAGA	840
Db	1183	AGCTGAAGAAGTCTTATGACAGAAATTAAGAACTGTATTTGATCATTTTGGAGAGCAAAAGAAAGA	1242
Qy	841	ACTTGAAAGGTATCAACAAAAATTTTGAAGAAGCTTAAAGAAATTTGGATTAAGAAAGCTAT	900
Db	1243	ACTTGAAAGGTATCAACAAAAATTTTGAAGAAGCTTAAAGAAATTTGGATTAAGAAAGCTAT	1302
Qy	901	CACGGCCACAAATTTGATATTTGTCGCCGCTTCTTATTAATCATACATCATATGCTCTGGC	960
Db	1303	TACAGCCAAATTTCTATAGTGTGCTGCTTCTCTGATCATATGATATTTATGCTCTGGC	1362
Qy	961	TTTCTGTGTATGGAGACCTCTGTGCTCTCCAGTAAATTTCTATTTGACAAGTACTCAC	1020
Db	1363	CTTCTGTGTATGGAGACACCTTGCTCTCTCAGAGGAAATTTCTATTTGACAAGTACTCAC	1422
Qy	1021	TGCTCTCTTTTCTGTATTAATTTAGGGGCTTTTATGTATTTGACACAGCATCCCAAGACTGA	1080
Db	1423	TGATTTCTTTTCTGTATTAATTTAGGGGCTTTTATGTATTTGACACAGCATCTCCAAGACTGA	1482
Qy	1081	AGCATTTTGCAAAAGCAAGAGGACACCTTATGAAATCTTCAAGATTAATTTGACATTAAGC	1140
Db	1483	AGCATTTTGCAAAAGCAAGAGGACACTTATGAAATCTTCAAGATTAATTTGATTAATTAAGC	1542
Qy	1141	AAGCATTTGACAGCTATTTGCAAGAGTGACATTAACCATTAATTTAAGAGGAAATTTGGA	1200
Db	1543	AAGATTTGACAGCTATTTGCAAGAGTGAGCCCAAAACCATTAATTTAAGGGAATTTTGGAA	1602
Qy	1201	ATTCAAAAATGTCTACTTTCAGTTAACCTTTCTCGAAAAGATTTAAGATCTTAAAGGCTCT	1260
Db	1603	ATTCAAAAATGTCTACTTTCAGTTAACCTTCCATTCGAAAAGATTTAAGATCTTGAAGGCTCT	1662
Qy	1261	CACCTGAAGTTGACAGAGTGGGACAGACATGGCGCTGTTTGGGAAACATGGCGCGGGGAA	1320
Db	1663	GACCTGAAGTTGACAGAGTGGGACAGAGTGGCGCTGTTTGGGAAACATGGCGCTGCGGAA	1722
Qy	1321	GAGCAGACCGTGCAGCTGATGACAGAGGCTTATGATCCCAAGATATGGCTGTGTAT	1380
Db	1723	GAGCAGACCGTGCAGCTGATGACAGAGGCTTATGATCCCAAGAGGAGTGGCTGATGT	1782
Qy	1381	TGATGGACAGGACATTAGGACCATTAATGTAAAGCATCTTGGGAAATTAAGTGTGGGT	1440
Db	1783	TGATGGACAGGATATTAGGACCATTAATGTAAAGCTTCTACGGGAAATTAAGTGTGGGT	1842
Qy	1441	GAGTCAAGAGCTGTGTGTTTGGCCACACAGATATGCTTAAACAAATTTGCTATGGCCCGGA	1500
Db	1843	GAGTCAAGAGCTGTATGTTTGGCCACACAGATATGCTTAAACAAATTTGCTATGGCCCGTGA	1902
Qy	1501	AAATGTACCATGATGATGATTAAGAAAGCTGTTAAGSAGCCAAATGCTATGATTTTAT	1560
Db	1903	AAATGTACCATGATGATGATTAAGAAAGCTTCTCAAGGAACCAATATGCTATGACTTTAT	1962
Qy	1561	CATGAATCTACTAATTAATTTGACACATCTGTGTTGAGAGAGAGGGGCCACGCTGAGTGG	1620
Db	1963	CATGAATCTGCCCATTAATTTGACACCTGTGTTGAGAGAGAGGGGCCAGATTGAGTGG	2022
Qy	1621	TGGACAGAAACAGAGATGGCCATTTGCTGGGGCCCTGTTGGCAACCCCAAGATTTCTCT	1680
Db	2023	TGGGCAACAGACAGAGATGGCCATTTGCTCAAGTGGCCCTGTTGGCAACCCCAAGATTTCTCT	2082

Oy	1681	GCTGATGAGGCAACGTCCAGCTCTGCACTGAAGTAAGTGAAGAGTGGTCCAGTGGCCCT	1740
Db	2083	GCTGATGAGGCCAACGTCAGCCTTGAGCACAGAAAGGAGGAGCTGGTTCAGGTGCTCT	2142
Oy	1741	GGATTAAGGCCAGAAAAGGCGGAGCTACCATTTGATAGTACATCGTTTGTCTCAACTTCG	1800
Db	2143	GGATTAAGGCCAGAAAAGGTGGAGCCACCATTGTGATAGCTCATTCGTTTGTCTCAACTTCG	2202
Oy	1801	TAAATCCGATGCTCATTTGCTGGTTTGTATGATGAGTCAATCTGTGAGAAAAGAAATCATCA	1860
Db	2203	TAAATCCGATGCTCATTCGCTGGTTTGTATGATAGTCAATCTGTGAGAAAAGAAATCATCA	2262
Oy	1861	TGAACCTCATGAAGAAGAGGCGCATTTACTTAAACTTGTGCATCATGTCAGACAGAGAGAA	1920
Db	2263	TGAACCTCATGAAGAAGAGGCGCATTTACTTAAACTTGTGCATCATGTCAGACAGAGAGAA	2322
Oy	1921	TGAAATTTGATTAGAAAAATGCCACTGTGTAATCCAAAAGTGAAGTATGCTCTTGGAAAT	1980
Db	2323	TGAAATTTGATTAGAAAAATGCAGCTGATGAATCCAAAAGTGAAGTATGATGCTCTTGGAAAT	2382
Oy	1981	GTCTCCAAAAGATTCAGGCTCCAGTTTAAATAAAAGAAAGATCACTGCGACAGAGATATACA	2040
Db	2383	GTCTCCAAAAGATTCAGGCTCCAGTCTAATAAAGAAAAGATCACTGCGATGAGTGTCCG	2442
Oy	2041	TGACACCAAGAGGCCAGACAGAGAAGCTTGTCAAAAGAGCATTTGAATGCAATGTACC	2100
Db	2443	TGATCTCAACAGCCCAAGACAGAGAAGCTTGTACCAAGAGGCTGTGGATGGAAGATATACC	2502
Oy	2101	TCCAGTTTCCTTCTGAGAGATCTCTAAGCTGAACCTCAACTGAATGGCTTTATTTTGTGGT	2160
Db	2503	TCCAGTTTCCTTCTTGGAGATTTGTGAAGCTAAATTTAACTGAATGGCTTTATTTTGTGGT	2562
Oy	2161	TGCTAATTTTGTGCTATTATTAACGAGAGGCGCTGCAACCGACATTTTCAATATTTTTC	2220
Db	2563	TGGTGATTTTGTGTGCCATTATTAATGGAAGGCGTGCACACGACATTTGCAATATTTTTC	2622
Oy	2221	AAGGATTAATAGGATCTTTACCCGAGATGAGATCCTGAAAACAAAACGACAAATAGTAA	2280
Db	2623	AAAGATTAATAGGAGTTTATACAAAGAAATGTATGATCCTGAAAACAAAACGACAAATAGTAA	2682
Oy	2281	CATGTTTTCTGATTTGTTTCTAGTCCCTTGGAAATTTTCTTTTAACTATTTTTCCTCCA	2340
Db	2683	CTGTGTTTCACTATTTGTTTCTAGGCCCTTGGAAATTTTCTTTTAACTATTTTTCCTCCA	2742
Oy	2341	GAGCTTCACATTTTGGCAAAAGCTGGGAGATCTTCACATAACGGGCTTCGATACATGGTTTT	2400
Db	2743	GAGTTTCAACATTTGGCAAAAGCTGGAGAGATCTTCACAAACGGGCTTCGATACATGGTTTT	2802
Oy	2401	CAGATCCATGCTGACAGAGAGATGTCAGCTGTTTGAATGAGCCCTAAAAACACCACTGGAGC	2460
Db	2803	CCGATCCATGCTGACACAGAGATGTGAGTTGTTTGAATGAGCCCTAAAAACACCACTGGAGC	2862
Oy	2461	ATTGACAACCAAGCTTTGCCAATGATGCGGCTCAAGTTTAAAGGGCTTATAGTTCACAGGCT	2520
Db	2863	ATTGACATACCAAGCTTCGCCAATGATGCTGCTCAAGTTTAAAGGGCTTATAGTTCACAGGCT	2922
Oy	2521	TGCTGTCATTTACCCAGAAATATGCAAAATCTTGGGACAGGCAATTTATATCCTTAATCTA	2580
Db	2923	TGCTGTAATTTACCCAGAAATATGCAAAATCTTGGGACAGGAAATTTATATCCTTCATCTA	2982
Oy	2581	TGCTTGGCAATTACACTTTTACTCTTATGACAAATGTATACCCATCATTTGCAATAGAGGAGGT	2640
Db	2983	TGCTTGGCAACTTACACTGTCTACTCTTATGACAAATTTTACCCATCATTTGCAATAGAGGAGGT	3042
Oy	2641	TGTTGAATGAATAATGTGTCTGTGGAACAACACTGAAAGATGAAGAAGAGCTAGAAGAGC	2700
Db	3043	TGTTGAATGAATAATGTGTCTGTGGAACAACACTGAAAGATGAAGAAGAGCTAGAAGAGC	3102
Oy	2701	TGGGAAGATTGCTACAGAAAGCCATCGAAAACCTTCGAACTGTTGTTTCTTTGACCTCGGGA	2760
Db	3103	TGGGAAGATGCTACTAGGAAGCAATTAAGAAAACCTTCGAAACCCGTTGTTTCTTTGACCTCAGGA	3162

OY 181 AATGGCTGCCATCATCATGAGCTGCACCTCCCTCATGATGCTGTTTGGAAACAT 240
|||||
Db 589 TTGGCTGCATCATCATCATGAGCTGCACCTCCCTCATGATGCTGTTTGGAAACAT 648
OY 241 GACACATAGCTTTGGCAATGACGAATTTGCAAGAAACAACTTTCCAGTTATATTTAA 300
|||||
Db 649 GACAGATATCTTTGCAAAATGACGAATTTGAGAGATCTGATGTCA-----AACATCAC 702
OY 301 TGAAGTATTTACGAACATATACAAACATTTGATCAACCATCTGGAGAGGAATGACAC 360
|||||
Db 703 TAATAGAGTATATCAATGATACAGGGTCTTCAATGATCTGAGAGAAACATGACAC 762
OY 361 GTATGCTATTTATACAGTGGAGTGCCTGGCTGCTGCTGCTGCTTACATCCAGGT 420
|||||
Db 763 GTATGCTATTTATACAGTGGAGTGTGCTGGCTGCTGCTGCTGCTTACATCCAGGT 822
OY 421 TTGATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 480
|||||
Db 823 TTGATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 882
OY 481 TGCATATGTCGACAGAGATTTGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 540
|||||
Db 883 TGCATATATGTCGACAGAGATTTGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 942
OY 541 CCGGCTCAGACAGAGATCTCTCCAAATTCATAGAGAAATTTGGCAAAATTTGGAATGTT 600
|||||
Db 943 CCGGCTCAGACAGAGATCTCTCTAAAGATTAAGAAATTTGGTGAACAAATTTGGAATGTT 1002
OY 601 CTTTACAGCAATATGACATTTTTCACCGGTTTATATAGTGGGCTTTACAGTGGTTGGA 660
|||||
Db 1003 CTTTACAGCAATATGACATTTTTCACCGGTTTATATAGTGGGCTTTACAGTGGTTGGA 1062
OY 661 GCTAACCCCTGTGATTTTGGCCATCACCCCTGTTTGGACCTTTCAGCCCATATGCGGC 720
|||||
Db 1063 GCTAACCCCTGTGATTTTGGCCATCACCCCTGTTTGGACCTTTCAGCCCATATGCGGC 1122
OY 721 AAGAGTATCTATCTTATCTATCTGATAAAGAACTTTGGCCTATGCAAAAGCTGGAGCA 780
|||||
Db 1123 AAGAGTATCTATCTTATCTATCTGATAAAGAACTTTCAGCTATGCAAAAGCTGGAGCA 1182
OY 781 AGCTGAAGAGTCTTTAGCAGCAATCAGACATGATGCTTGGAGAGCAAAAGAAAGA 840
|||||
Db 1183 AGCTGAAGAGTCTTTAGCAGCAATCAGACATGATGCTTGGAGAGCAAAAGAAAGA 1242
OY 841 ACCTGAAGAGTCTCAACAAATTTAGAGAAAGCTAAAGAAATGGGCTAAAGAAAGCTAT 900
|||||
Db 1243 ACCTGAAGAGTCTCAACAAATTTAGAGAAAGCTAAAGAAATGGGCTAAAGAAAGCTAT 1302
OY 901 CACGGCCACATTTCTATTTGTCGCTTCTATGATGATCATCATATGCTGTGGC 960
|||||
Db 1303 TACAGCCAAATTTCTATTTAGTGGTGTGCTGCTGCTGATCATCTTATGCTGTGGC 1362
OY 961 TTTCTGATGAGGACCTCCCTGCTCCTCCAGTGAATTTCTATTTGGACAAAGTACAC 1020
|||||
Db 1363 CTCTGATGAGGACCACTGCTGCTCCTCAGGGGAAATTTCTATTTGGACAAAGTACAC 1422
OY 1021 TGTCTCTCTCTCTCTATTTATTTGGGCTTTTATGATGAGAGGATCCCAACATTTGA 1080
|||||
Db 1423 TGTCTCTCTCTCTCTATTTATTTGGGCTTTTATGATGAGAGGATCCCAACATTTGA 1482
OY 1081 AGCATTTGCAAAAGCAGAGAGAGAGCTTATGAAATCTTCAAGATTAATTTGACATTAAC 1140
|||||
Db 1483 AGCATTTGCAAAATGCAAGAGAGAGAGCTTATGAAATCTTCAAGATTAATTTGATTAAC 1542
OY 1141 AAGCATTTGACAGTATTTGAGAGAGTGAACATTAACAGATTAATTTAAGAGAAATTTGGA 1200
|||||
Db 1543 AAGTATTTGACAGTATTTGAGAGAGTGAACATTAACAGATTAATTTAAGAGAAATTTGGA 1602
OY 1201 ATTGAAATTTGACATTTGACATTTCCCTCTCGAAAAAGATTTAAGATTTAAGGGCT 1260
|||||
Db 1603 ATTGAAATTTGACATTTGACATTTCCCTCTCGAAAAAGATTTAAGATTTAAGGGCT 1662
OY 1261 CAACCTGAAGTTGACAGTGGGAGACAGTGGGCTGTTGGGAACAGTGGCTGGGGGA 1320
|||||

Db 1663 GAACCTGAAGTTGACAGTGGGAGACAGTGGGCTGTTGGAAACAGTGGCTGGGAA 1722
|||||
OY 1321 GACACAGAGCTGTCAGAGTATGACAGAGCTTATGACCCCAAGATGATGCTGTAT 1380
|||||
Db 1723 GACACAGAGCTGTCAGAGTATGACAGAGCTTATGACCCCAAGATGATGCTGTAT 1782
OY 1381 TGAATGACAGAGATTTAGAGAGTATGATTAAGATCTTTGGGAAATTTCTGCTGTGT 1440
|||||
Db 1783 TGAATGACAGAGATTTAGAGAGTATGATTAAGATCTTTAGAGAAATTCATGCTGTGT 1842
OY 1441 GAGTCAGAGAGCTGTTGTTGGCCACAGATATGATGATGATGATGATGATGATGAT 1500
|||||
Db 1843 GAGTCAGAGAGCTGTTATGTTTTCACACAGATGATGATGATGATGATGATGATGAT 1902
OY 1501 AATGTCACCATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1560
|||||
Db 1903 AATGTCACCATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1962
OY 1561 CATGAACTACCTATTAATTTGACACTGCTGTTGGAGAGAGAGGGGCCACAGATCTTCT 1680
|||||
Db 1963 CATGAACTACCTATTAATTTGACACTGCTGTTGGAGAGAGAGGGGCCACAGATCTTCT 2022
OY 1621 TGCACAGAAACAGAGATTCGCAATGCTGCGGCCCTGCTGCTGCTGCTGCTGCTGCT 1680
|||||
Db 2023 TGCACAGAAACAGAGATTCGCAATGCTGCGGCCCTGCTGCTGCTGCTGCTGCTGCT 2082
OY 1681 GCTGATGAGCAAGCAGTACGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1740
|||||
Db 2083 GCTGATGAGCAGCAGTACGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2142
OY 1741 GATTAAGGCAAGAAAGGCGGCTACCTACATTTGATGATGATGATGATGATGATGAT 1800
|||||
Db 2143 GATTAAGGCAAGAAAGGCGGCTACCTACATTTGATGATGATGATGATGATGATGAT 2202
OY 1801 TAATGCGCATGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1860
|||||
Db 2203 TAATGCGCATGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2262
OY 1861 TGAATCATGAAAGAGAGAGGCTATTTCAAACTTGTACAAATGACAGAGAGAA 1920
|||||
Db 2263 TGAATCATGAAAGAGAGAGGCTATTTCAAACTTGTCAATGACAGAGAGAA 2222
OY 1921 TGAATGAGTGAAGAAATGCACTGATGATGATGATGATGATGATGATGATGATGAT 1980
|||||
Db 2223 TGAATGAGTGAAGAAATGCACTGATGATGATGATGATGATGATGATGATGATGAT 2282
OY 1981 GTCTCCAAAGATTTCAAGGCTCAGTTATTAAGAAAGATCAATGCAAGATGATGAT 2040
|||||
Db 2383 GTCTCCAAAGATTTCAAGGCTCAGTTATTAAGAAAGATCAATGCAAGATGATGAT 2442
OY 2041 TGCACCAAGAGGCAAGAGAGAGGCTGATGATGATGATGATGATGATGATGATGAT 2100
|||||
Db 2443 TGCATCAAGAGGCAAGAGAGAGGCTGATGATGATGATGATGATGATGATGATGAT 2502
OY 2101 TCCAGTTTCTCTGAGAGATCTGAAGCTGAACCTCAATGAAATGATGATGATGATGAT 2160
|||||
Db 2503 TCCAGTTTCTCTGAGAGATCTGAAGCTGAACCTCAATGAAATGATGATGATGATGAT 2562
OY 2161 TGTATATTTTGTCTATTAAGAGAGGCTGCAACAGCATTTTCAATTAATTTTC 2220
|||||
Db 2563 TGTATATTTTGTCTATTAAGAGAGGCTGCAACAGCATTTTCAATTAATTTTC 2622
OY 2221 AAGATTTATAGGATTTTACCAAGAAATGATGATGATGATGATGATGATGATGAT 2280
|||||
Db 2623 AAGATTTATAGGATTTTACCAAGAAATGATGATGATGATGATGATGATGATGAT 2682
OY 2281 CATGTTTCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2340
|||||
Db 2683 CATGTTTCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2742
OY 2341 GGGCTTCACATTTGCAAGAGCTGGGAGATCTCAGTACGGCTTGCATACATGCTTT 2400
|||||

Db 2743 GGGTTTCACATTGGCAAGCTGGAGAGATCTCCACCAAGGCGCTCCGATACATGGTTTT 2802
 QY 2401 CAGATCCATGCTGAGACAGATGTCAGTGGTTTGATGAGCCCTTAATAAACCACTGGAGC 2460
 Db 2803 CCGATCCATGCTGAGACAGATGTCAGTGGTTTGATGAGCCCTTAATAAACCACTGGAGC 2862
 QY 2461 ATTGACCAACGCGTTGCCAATGATGGCGCTCAAGTTAAAGGGGCTATAGTCCAGGCT 2520
 Db 2863 ATTGACTAACCGGCTCCGCAATGATGCTGCTCAAGTTAAAGGGGCTATAGTCCAGGCT 2922
 QY 2521 TGTGTAATTAATCCAGAAATATAGAAATCTTGGGACAGGCGATTTATATCTTAACTTA 2580
 Db 2923 TGTGTAATTAATCCAGAAATATAGAAATCTTGGGACAGGCGATTTATATCTTAACTTA 2982
 QY 2581 TGTGTAATTAATTAACACTTTTACTCTTACCAATTTGATACCATATGCAATGCAATGCA 2640
 Db 2983 TGTGTAATTAATTAACACTTTTACTCTTACCAATTTGATACCATATGCAATGCAATGCA 3042
 QY 2641 TGTGTAATTAATTAATGTTGCTGTCGACACGACTGAAAGATTAAGAAAGAGCTAGAAGAG 2700
 Db 3043 TGTGTAATTAATTAATGTTGCTGTCGACACGACTGAAAGATTAAGAAAGAGCTAGAAGAG 3102
 QY 2701 TGGGAAGATTGCTACAGAAAGCCATCGAAAACTTGGAACTGTTGTTCTTTGACTCGGGA 2760
 Db 3103 TGGGAAGATTGCTACAGAAAGCCATCGAAAACTTGGAACTGTTGTTCTTTGACTCGGGA 3162
 QY 2761 GCAGAAATTGAATACATGATATGACACAGATTTGCAAGTACCATACAAACTCTTTGAG 2820
 Db 3163 GCAGAAATTGAATACATGATATGACACAGATTTGCAAGTACCATACAAACTCTTTGAG 3222
 QY 2821 GAAAGCACATCTTCGGGGGCTCATTTTCTATACACCCAGCAATGATGATATTTTCTTA 2880
 Db 3223 GAAAGCACATCTTCGGGGGCTCATTTTCTATACACCCAGCAATGATGATATTTTCTTA 3282
 QY 2881 TGTGCTGTTTCCGCTTGGTGGCTTGTGCTTGTGCTTGTGCTTGTGCTTGTGCTTGTGCT 2940
 Db 3283 TGTGCTGTTTCCGCTTGGTGGCTTGTGCTTGTGCTTGTGCTTGTGCTTGTGCTTGTGCT 3342
 QY 2941 TGTGCTGTTTGGTATCTGACATGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTT 3000
 Db 3343 TGTGCTGTTTGGTATCTGACATGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTT 3402
 QY 3001 ATTGCTCTGACTATGCAAGCAAGCAAGTATGACAGCCCGCTCATGATGATGATGATGAT 3060
 Db 3403 ATTGCTCTGACTATGCAAGCAAGCAAGTATGACAGCCCGCTCATGATGATGATGATGAT 3462
 QY 3061 AAAAAGCCCTGATTTGACACTACAGCCCTACAGGCGCTTAAGCCAAATACGTTGGAAG 3120
 Db 3463 AAAAAGCCCTGATTTGACACTACAGCCCTACAGGCGCTTAAGCCAAATACGTTGGAAG 3522
 QY 3121 AAATGTGACATTTATGAGTGTGTTCAACTATCCCACTGACAGACATCCCGTGT 3180
 Db 3523 AAATGTGACATTTGTTGTAATGTTGTAATCACTATCCCACTGACAGACATCCCGTGT 3582
 QY 3181 CCAGGGGCTGAGCTGAGGTTGAAGAGGCGCAGACGCTGGCCCTGTAGTGAAGAGTGG 3240
 Db 3583 TCAGGAGACTAGGCTGAGGTTGAAGAGGCGCAGACGCTGGCCCTGTAGTGAAGAGTGG 3642
 QY 3241 CTGTGGAAGAGCAGATTTGCTAGCTCTAGAGCGCTCTTAATGAGCCCTTGGCTGTTC 3300
 Db 3643 CTGTGGAAGAGCAGATTTGCTAGCTCTAGAGCGCTCTTAATGAGCCCTTGGCTGTTC 3702
 QY 3301 AGTGTAAATTTGATGCAAGAAAGATTAAGCACTTAATGCTGAGTCCGAGCACACT 3360
 Db 3703 AGTGTAAATTTGATGCAAGAAAGATTAAGCACTTAATGCTGAGTCCGAGCACACT 3762
 QY 3361 GGGCATCTGTTCTAGAGAGCCCATCTGTTTACGAGCAATGCGGAGAACTTGCCTTA 3420
 Db 3763 GGGCATCTGTTCTAGAGAGCCCATCTGTTTACGAGCAATGCGGAGAACTTGCCTTA 3822
 QY 3421 TGGAGCAACACCGGGTGTATCATAGAAAGATTTATGCAAGGAGCAGCAAGAGGCAA 3480
 Db 3823 TGGAGCAACACCGGGTGTATCATAGAAAGATTTATGCAAGGAGCAGCAAGAGGCAA 3882

QY 3481 CATACACACTTCATCCAGACACTCTCTGAGAAATACACACAGAGTAGAGACAAAG 3540
 Db 3883 CATACACTCTCTCATGCTAGCTACGCTTAATTAATATAGCACTAAAGTAGAGACAAAG 3942
 QY 3541 AACCAGCTCTCTGTTGAGGAGAAACAGGCGCATTTGCTAGTGGCTCTTGTAGAA 3600
 Db 3943 AACTCAGCTCTGTTGAGGAGAAACAGGCGCATTTGCTAGTGGCTCTTGTAGAA 4002
 QY 3601 GCCCATATTTTGTGTTTGTGATGAGCTACATGCTCTGATATACAGAAAGTAAAGST 3660
 Db 4003 GCCCATATTTTGTGTTTGTGATGAGCTACATGCTCTGATATACAGAAAGTAAAGST 4062
 QY 3661 TGTCCAAAGAGCCCTGAGCAACCAAGAGAGGCGCCACTGCTGATTTGTATGCCACCG 3720
 Db 4063 TGTCCAAAGAGCCCTGAGCAACCAAGAGAGGCGCCACTGCTGATTTGTATGCCACCG 4122
 QY 3721 CTGTCCACCATCCAGAAATGACATTTTAATAGTGTGTTTCAAGATGGCAAGTCAAG 3780
 Db 4123 CTGTCCACCATCCAGAAATGACATTTTAATAGTGTGTTTCAAGATGGCAAGTCAAG 4182
 QY 3781 GCATGCAACATCAACAGCTCTGCGCCAGAAAGCATTTTCTCATGTCAGTGT 3840
 Db 4183 GCATGCAACATCAACAGCTCTGCGCCAGAAAGCATTTTCTCATGTCAGTGT 4242
 QY 3841 CCAGGCTGAGCAAAAGCGCTAGTGAATGTGGCCATATGAGCTGTAAATTTTAAAT 3900
 Db 4243 CCAGGCTGAGCAAAAGCGCCAGTGAATGTGGCCATATGAGCTGTAAATTTTAAAT 4302
 QY 3901 ATTGTTGTTAAACATGTCATTTATCAAAAGTAAAGTGAAGCACTTACGAAACT 3960
 Db 4303 ATT--TGTTAATATGACATTTATCAAAAGTAA--AGCAAACTTACGAAAT 4355
 QY 3961 ATGTGAACACTGTTTAACTTCTGCAACTGAGATGATATCCACCAAGTTGAG 4020
 Db 4356 ATGTGAACACTGTTTAACTTCTGCAACTGAGATGATATCCACCAAGTTGAG 4403
 QY 4021 AACTTCAGATTTTAAATTAATTAAGAAACAA--AAGAACTTATCTGATGAATAAAT 4078
 Db 4404 ----AGAGACTGTAAATTAAGAAAGAGTGAAGATCATCAAGTGAGAGAAAT 4458
 QY 4079 ACTGCTGTTAAATGCAATTAATTAATTAAGATTAATCAAGATGAATTTGTTAAAT 4138
 Db 4459 CATAGTTAAATGCAATTAATTAATTAAGATTAATCAAGATGAATTTGTTAAAT 4518
 QY 4139 TGTAAATTTTGTATTAATTTT-----ATTGTAATCTGCTTGTGGAAGAT 4190
 Db 4519 ATGTGTAATTTGTTAAATTTTCCATTTTGAATGACTGTAATGACTGCTGCTGAAGAT 4578
 QY 4191 TATGAAGTGTAAAGTACTG--AATGTTGAATTAAGTGTAGTATATTAAGTAA 4249
 Db 4579 TATGAAGTGTAAAGTATTAAGTATTTGATTAAGTGT--TGTAAATTAAGTAA 4635
 QY 4250 ACTTTTAAAT 4259
 Db 4636 ACTTTTAAAT 4645

Search completed: December 9, 2002, 21:10:40
 Job time : 1403 secs

GenCore version 5.1.3
Copyright (c) 1993 - 2002 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: December 9, 2002, 16:58:52 ; Search time 993 Seconds
(without alignments)
9704.230 Million cell updates/sec

Title: US-09-672-725C-1

Perfect score: 4279

Sequence: 1 ggaagcgcgaagtcgcatg.....caaaaaaaaaaaaaaa 4279

Scoring table:

IDENTITY NUC
Gapop 10.0 , Gapext 1.0

Searched: 2185239 seqs, 1125999159 residues

Total number of hits satisfying chosen parameters: 4370478

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

N Geneseq 101002: *
1: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1980.DAT: *
2: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1981.DAT: *
3: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1982.DAT: *
4: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1983.DAT: *
5: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1984.DAT: *
6: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1985.DAT: *
7: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1986.DAT: *
8: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1987.DAT: *
9: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1988.DAT: *
10: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1989.DAT: *
11: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1990.DAT: *
12: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1991.DAT: *
13: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1992.DAT: *
14: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1993.DAT: *
15: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1994.DAT: *
16: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1995.DAT: *
17: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1996.DAT: *
18: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1997.DAT: *
19: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1998.DAT: *
20: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1999.DAT: *
21: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA2000.DAT: *
22: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA2001A.DAT: *
23: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA2001B.DAT: *
24: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA2002.DAT: *

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	4279	100.0	4279	22	AAD03488
2	4277.4	100.0	4279	22	AAD03504
3	4275.8	99.9	4279	22	AAD03505
4	4272.6	99.9	4279	22	AAD03506
5	4221.8	98.7	4317	22	AAD03489
6	3341.2	78.1	4669	8	AAAT0752
7	3335	77.9	4669	14	AA052726
8	3334.8	77.9	4378	11	AA004522
9	3333.2	77.9	4669	19	AAV32645

10	3333.2	77.9	4669	24	ABK52041	cDNA encoding huma
11	3324.4	77.7	4646	21	AAZ94738	Human ATP binding
12	3324.4	77.7	4646	21	AAD36994	Human mdr1 gene.
13	3324.4	77.7	4646	24	ABL68592	Kidney cancer rela
14	3324.4	77.7	4646	24	ABL68880	Human cancer rela
15	3322.6	77.6	4646	22	AAH57442	Human intestine ce
16	3319.6	77.6	4646	22	AAO72872	Human multidrug re
17	3311.8	77.4	4186	22	AAE6127	Cynomolgous monke
18	3311.2	77.4	4195	22	AAE6128	Mutated human P-g1
19	3270.6	76.4	4264	19	AAV66533	Human MDR-1 DNA.
20	3270.6	75.9	4264	19	AAV66534	Hybrid vector pSF-
21	3247.6	75.9	3988	21	AAZ88973	Retroviral M4 mdr-
22	3239.6	75.7	6505	17	AAZ13394	Human wild-type mu
23	3226.6	75.4	8630	21	AAZ24042	Human BCRP DNA rel
24	3223.4	75.3	8630	21	AAZ24041	Human BCRP DNA rel
25	3202	74.8	3860	21	AAZ49332	Human G185V mutant
26	3202	74.8	3860	21	ABA94365	Human BCRP DNA rel
27	3200.4	74.8	3860	21	AAZ49333	Human BCRP DNA rel
28	3200.4	74.8	3860	24	ABA94366	Human BCRP DNA rel
29	3190.2	74.6	3840	24	ABL91687	Human BCRP DNA rel
30	2936.2	68.6	4788	21	AAZ49335	Human BCRP DNA rel
31	2936.2	68.6	4788	24	ABA94368	Human BCRP DNA rel
32	2921.4	68.3	4425	21	AAZ52048	Murine multidrug re
33	2817	68.2	4369	21	AAZ52047	Murine multidrug re
34	2730.2	63.8	4313	14	AAQ38950	Rat multidrug resi
35	2716.6	63.5	4189	21	AAZ49334	Murine multidrug re
36	2716.6	63.5	4189	24	ABA94367	Murine multidrug re
37	2705	63.2	4233	21	AAZ90198	Mouse BCRP DNA rel
38	2705	63.2	4233	22	AAZ27498	Mouse BCRP DNA rel
39	2677.4	62.6	4254	24	ABK63517	Rat mdr1b2 multidr
40	2240	52.3	3924	21	AAZ94742	Rat mdr1b2 multidr
41	2240	52.3	3924	21	AAZ88974	Rat mdr1b2 multidr
42	2240	52.3	3924	24	ABN95801	Human ATP binding
43	2118	49.3	2726	15	AAQ70907	Gene #2299 used to
44	2118	49.5	2726	15	AAQ70916	Multidrug-resistan
45	2118	49.5	2726	18	AAZ43322	Multidrug-resistan

ALIGNMENTS

RESULT 1	
ID	AAD03488
AAID03488	standard; cDNA, 4279 BP.
AC	AAID03488:
DT	13-JUN-2001 (first entry)
DE	Dog P-glycoprotein (PGP) cDNA #1.
KW	Dog; P-glycoprotein; PGP; multidrug transporter; MDRL;
KW	drug bioavailability; transgenic animal; genetic model; ss.
OS	Canis familiaris.
XX	
XX	Key
FT	Location/Qualifiers
FT	17..3862
FT	/tag= a
FT	/product= "Dog P-glycoprotein (PGP) #1"
PN	WO200123540-A2.
XX	
PD	05-APR-2001.
XX	
PF	28-SEP-2000; 2000WO-US26767.
XX	
PR	28-SEP-1999; 99US-0156510.
XX	
PA	(GENT-) GENTEST CORP.
XX	
PI	Stocker PJ, Steimel-crespi DT, Crespi CL, Reif TC, Patten CJ.
XX	

DR MPI: 2001-235373/24.
DR P-PSDB: AAE00303.
XX New dog P-glycoproteins (Pgp) and their encoding nucleic acids, useful
PT for determining the bioavailability of drugs and for screening for dog
PT Pgp inhibitors -
XX
PS Claim 3: Page 58-63; 111pp; English.
XX The invention relates to dog P-glycoprotein (Pgp) also referred
CC as multidrug transporter (MDR1) and nucleic acids encoding them.
CC The invention also includes fragments and biologically functional
CC variants of dog P-glycoprotein. Pgp and their nucleic acids are
CC useful for determining the bioavailability of drugs and for
CC screening Pgp inhibitors. They are useful for the diagnosis and
CC treatment of conditions characterised by Pgp activity, by
CC reducing or increasing Pgp activity in a cell. Pgp nucleic acids
CC are used as oligonucleotide probes. Complements of Pgp nucleic
CC acids are useful as antisense oligonucleotides, to induce a Pgp
CC 'knockout' phenotype. They are used to prepare a non-human
CC transgenic animal, which are valuable as genetic models for
CC human diseases.
CC The present sequence is dog P-glycoprotein (Pgp) cDNA. This
CC sequence is also referred as Genotype C cDNA. The
CC Pgp enzyme functions as an efflux pump exporting small molecules
CC across the cell membrane. This enzyme is a member of the ABC
CC transporter family.
SQ

Sequence 4279 BP: 1294 A; 834 C; 1008 G; 1143 T; 0 other;

Query Match 100.0%; Score 4279; DB 22; Length 4279;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 4279; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGAGCGCCGAGGTGCGGATGATCTCGAAGGAGGCGCTAAGGAGTGCAGAGAAACTT 60
DB 1 GGAGCGCGAGGTGCGGATGATCTCGAAGGAGGCGCTAAGGAGTGCAGAGAAACTT 60
QY 61 CTGGAATAATGGCCAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 120
DB 61 CTGGAATAATGGCCAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 120
QY 121 CACGTTTGCATGTTGCTGATTTCAAAATGGCTTATGATGTTGTTGGGAC 180
DB 121 CACGTTTGCATGTTGCTGATTTCAAAATGGCTTATGATGTTGTTGGGAC 180
QY 121 CACGTTTGCATGTTGCTGATTTCAAAATGGCTTATGATGTTGTTGGGAC 180
DB 121 CACGTTTGCATGTTGCTGATTTCAAAATGGCTTATGATGTTGTTGGGAC 180
QY 181 AATGGCTGCATCATCATGAGCTGACCTCCCTCATGATGCTGTTTGGAAAT 240
DB 181 AATGGCTGCATCATCATGAGCTGACCTCCCTCATGATGCTGTTTGGAAAT 240
QY 181 AATGGCTGCATCATCATGAGCTGACCTCCCTCATGATGCTGTTTGGAAAT 240
DB 181 AATGGCTGCATCATCATGAGCTGACCTCCCTCATGATGCTGTTTGGAAAT 240
QY 241 GACAGATAGCTTTGCAATTCAGAGATTTCCAGAAACAACTTTCCAGTATATTA 300
DB 241 GACAGATAGCTTTGCAATTCAGAGATTTCCAGAAACAACTTTCCAGTATATTA 300
QY 301 TGAAGATATTACGAACAATACACAATTTCAATCAACACTGAGAGAGAAATGACC 360
DB 301 TGAAGATATTACGAACAATACACAATTTCAATCAACACTGAGAGAGAAATGACC 360
QY 361 GTATGCTATTATTACAGTGGATCGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCT 420
DB 361 GTATGCTATTATTACAGTGGATCGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCT 420
QY 421 TTTCAATTCGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 480
DB 421 TTTCAATTCGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 480
QY 481 TGTATATTCGACAGAGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 540
DB 481 TGTATATTCGACAGAGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 540
QY 541 CCGGCTCAGACAGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 600
DB 541 CCGGCTCAGACAGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 600

QY 601 CTTTCACTAATAGCAACAATTTTTCACCGGTTTATAGTGGGTTTACACGTGTTGAA 660
DB 601 CTTTCACTAATAGCAACAATTTTTCACCGGTTTATAGTGGGTTTACACGTGTTGAA 660
QY 661 GCTAACCTTGTATTTTGGCCATCAGCCCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 720
DB 661 GCTAACCTTGTATTTTGGCCATCAGCCCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 720
QY 721 AAGATATCTATCTTCAATTTACATTAAGACCTTGGCTTATGAGAAAGCTGAGAC 780
DB 721 AAGATATCTATCTTCAATTTACATTAAGACCTTGGCTTATGAGAAAGCTGAGAC 780
QY 781 AGCTGAAGAGCTTGTGACAGCAATTCAGAACTGATGCTGCTTGGAGCAAAAGAA 840
DB 781 AGCTGAAGAGCTTGTGACAGCAATTCAGAACTGATGCTGCTTGGAGCAAAAGAA 840
QY 841 ACTTGAAGGTACACAAAATTTAGAAAGCTTAAGAAATTTGGATTAAGAAAGCTAT 900
DB 841 ACTTGAAGGTACACAAAATTTAGAAAGCTTAAGAAATTTGGATTAAGAAAGCTAT 900
QY 901 CAGGCGCAACATTTCTATGTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 960
DB 901 CAGGCGCAACATTTCTATGTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 960
QY 961 TTTTGTATGAGGACCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1020
DB 961 TTTTGTATGAGGACCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1020
QY 1021 TGTCTTCTTTCTGATTTAATTTGGGCTTTATGATGATGAGACAGCATATGCTGCT 1080
DB 1021 TGTCTTCTTTCTGATTTAATTTGGGCTTTATGATGATGAGACAGCATATGCTGCT 1080
QY 1081 ACATTTTCAAAAGCAAGAGAGAGAGCTTATGAATTTTCAAGATTAATTTCAATTA 1140
DB 1081 ACATTTTCAAAAGCAAGAGAGAGAGCTTATGAATTTTCAAGATTAATTTCAATTA 1140
QY 1141 AAGCATGACAGCTTTTGAAGAGTGGACATTAACCAAGATTAATTTAAGGAAATTTGA 1200
DB 1141 AAGCATGACAGCTTTTGAAGAGTGGACATTAACCAAGATTAATTTAAGGAAATTTGA 1200
QY 1201 ATTCAAAATGCTACTTCACTTCACTTCACTTCACTTCACTTCACTTCACTTCACT 1260
DB 1201 ATTCAAAATGCTACTTCACTTCACTTCACTTCACTTCACTTCACTTCACTTCACT 1260
QY 1261 CAACCTGAAGCTTCAAGAGTGGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1320
DB 1261 CAACCTGAAGCTTCAAGAGTGGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1320
QY 1321 GAGCAGACCGTGCAGCTGATGACAGAGGCTTATGACCCACAGATGCTGCTGAT 1380
DB 1321 GAGCAGACCGTGCAGCTGATGACAGAGGCTTATGACCCACAGATGCTGCTGAT 1380
QY 1381 TGAATGACAGGATTAAGGACATTAATTAAGGACATTTCCGGAATTTACTGCTG 1440
DB 1381 TGAATGACAGGATTAAGGACATTAATTAAGGACATTTCCGGAATTTACTGCTG 1440
QY 1441 GAGTCAGAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1500
DB 1441 GAGTCAGAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1500
QY 1501 AATGTCACCATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1560
DB 1501 AATGTCACCATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1560
QY 1561 CATGAACCTACTATTAATTTGACACTTGTGTTGAGAGAGAGAGAGAGAGAGAGAG 1620
DB 1561 CATGAACCTACTATTAATTTGACACTTGTGTTGAGAGAGAGAGAGAGAGAGAGAG 1620
QY 1621 TGGACAGAAACAGAAATGCGCATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1680
DB 1621 TGGACAGAAACAGAAATGCGCATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1680


```
QY 1681 GCTGGATGAGGCAACCTGACTCTGTGACACTGAAGTGAGAGCGTTCAGGTGCCCCCT 1740
DB 1681 GCTGGATGAGGCAACCTGACTCTGTGACACTGAAGTGAGAGCGTTCAGGTGCCCCCT 1740
QY 1741 GGAATAGGCGCAAAAAGGCCGAGCTACCAATTTGTATAGCTCATGTTTGTCTACAGTTG 1800
DB 1741 GGAATAGGCGCAAAAAGGCCGAGCTACCAATTTGTATAGCTCATGTTTGTCTACAGTTG 1800
QY 1801 TAATGCCGATGCTATGCTGTTGATGATGAGTCACTATTGTGAGAGAAATCATGA 1860
DB 1801 TAATGCCGATGCTATGCTGTTGATGATGAGTCACTATTGTGAGAGAAATCATGA 1860
QY 1861 TGAAGTCAATGAAAGAGAGAGGSCATTTACTTCAAACTGTGACAAATGAGACAGAGAGAA 1920
DB 1861 TGAAGTCAATGAAAGAGAGAGGSCATTTACTTCAAACTGTGACAAATGAGACAGAGAGAA 1920
QY 1921 TGAATTTGAGTTAGAAAATGCGCACTGSGTAATCCAAAAGTGAAGATGATGCTTGGAAAT 1980
DB 1921 TGAATTTGAGTTAGAAAATGCGCACTGSGTAATCCAAAAGTGAAGATGATGCTTGGAAAT 1980
QY 1981 GTCGCCAAAAGATTCAGGGTCCAGTTTAAATAAAAGAGATCAACTGCGAGAGATATCA 2040
DB 1981 GTCGCCAAAAGATTCAGGGTCCAGTTTAAATAAAAGAGATCAACTGCGAGAGATATCA 2040
QY 2041 TGCACCAAGGCCAAGACAGAGAAAGCTGGTACAAAAGAGACTTGAATGAGAAATGTACC 2100
DB 2041 TGCACCAAGGCCAAGACAGAGAAAGCTGGTACAAAAGAGACTTGAATGAGAAATGTACC 2100
QY 2101 TCCAGTTCTCTTGTGAGAGATTCGAAAGCTGAACTCAATGAAATGSCCTTATTTTGTGCT 2160
DB 2101 TCCAGTTCTCTTGTGAGAGATTCGAAAGCTGAACTCAATGAAATGSCCTTATTTTGTGCT 2160
QY 2161 TGGTATATTTTGTGATATTAAGAGAGGCGCTGCAACGACATTTCAATATATTTTC 2220
DB 2161 TGGTATATTTTGTGATATTAAGAGAGGCGCTGCAACGACATTTCAATATATTTTC 2220
QY 2221 AAGGATTAAGGAGATCTTACCCGAGATGAGAGATCTGAAACAAAAGACAGCAATAGTAA 2280
DB 2221 AAGGATTAAGGAGATCTTACCCGAGATGAGAGATCTGAAACAAAAGACAGCAATAGTAA 2280
QY 2281 CATGTTTCTGATTTGTTTCTTACTGTTGAAATATTTCTTTTATCATTTTCCCTCA 2340
DB 2281 CATGTTTCTGATTTGTTTCTTACTGTTGAAATATTTCTTTTATCATTTTCCCTCA 2340
QY 2341 GGGCTTCACATTTGGCAAGCTGGGGAGATCTCTCAATGAGGCGCTGATACATGTTTTT 2400
DB 2341 GGGCTTCACATTTGGCAAGCTGGGGAGATCTCTCAATGAGGCGCTGATACATGTTTTT 2400
QY 2401 CAGATCCATGCTGAGACAGATGTCAGCTGTTTGAATGACCTTAAAAACACCACTGGAGC 2460
DB 2401 CAGATCCATGCTGAGACAGATGTCAGCTGTTTGAATGACCTTAAAAACACCACTGGAGC 2460
QY 2461 ATTGACAACAGGCTTCCCAATGATGAGGCTCAAGTTAAAGGGGCTATAGTTCCAGGCT 2520
DB 2461 ATTGACAACAGGCTTCCCAATGATGAGGCTCAAGTTAAAGGGGCTATAGTTCCAGGCT 2520
QY 2521 TGTGTCATTAACCCAGATATTAACAATCTTGGACAGGCAATATATATCTTATATCA 2580
DB 2521 TGTGTCATTAACCCAGATATTAACAATCTTGGACAGGCAATATATATCTTATATCA 2580
QY 2581 TGTGTCATTAACCCAGATATTAACAATCTTGGACAGGCAATATATATCTTATATCA 2580
DB 2581 TGTGTCATTAACCCAGATATTAACAATCTTGGACAGGCAATATATATCTTATATCA 2580
QY 2641 TGTGTCATTAACCCAGATATTAACAATCTTGGACAGGCAATATATATCTTATATCA 2640
DB 2641 TGTGTCATTAACCCAGATATTAACAATCTTGGACAGGCAATATATATCTTATATCA 2640
QY 2701 TGGGAAGATTTGCTACAGAGGCAATGAAAGCTTCCGAATGTTGTTTCTTGTACTCGGGA 2760
DB 2701 TGGGAAGATTTGCTACAGAGGCAATGAAAGCTTCCGAATGTTGTTTCTTGTACTCGGGA 2760
QY 2761 GAGAGATTTGAAATACATGATATGACAGAGTTTGCAGTACATACAGAAACCTTTGAG 2820
DB 2761 GAGAGATTTGAAATACATGATATGACAGAGTTTGCAGTACATACAGAAACCTTTGAG 2820
DB 2761 GAGAGATTTGAAATACATGATATGACAGAGTTTGCAGTACATACAGAAACCTTTGAG 2820
QY 2821 GAAAGCACACATCTTGGGGGTCTCATTTTATACCCAGGCAATGATATTTTTCCTA 2880
DB 2821 GAAAGCACACATCTTGGGGGTCTCATTTTATACCCAGGCAATGATATTTTTCCTA 2880
QY 2881 TGTGTCATTAACCCAGATATTAACAATCTTGGACAGGCAATATATATCTTATATCA 2940
DB 2881 TGTGTCATTAACCCAGATATTAACAATCTTGGACAGGCAATATATATCTTATATCA 2940
QY 2941 TGTGTCATTAACCCAGATATTAACAATCTTGGACAGGCAATATATATCTTATATCA 2940
DB 2941 TGTGTCATTAACCCAGATATTAACAATCTTGGACAGGCAATATATATCTTATATCA 2940
QY 3001 ATTGTCATTAACCCAGATATTAACAATCTTGGACAGGCAATATATATCTTATATCA 3060
DB 3001 ATTGTCATTAACCCAGATATTAACAATCTTGGACAGGCAATATATATCTTATATCA 3060
QY 3061 AAAAGGCCCTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 3120
DB 3061 AAAAGGCCCTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 3120
QY 3121 AAATGTCATTAACCCAGATATTAACAATCTTGGACAGGCAATATATATCTTATATCA 3180
DB 3121 AAATGTCATTAACCCAGATATTAACAATCTTGGACAGGCAATATATATCTTATATCA 3180
QY 3181 CCAAGGGCTGAGCTGAGGCTGAGGCTGAGGCTGAGGCTGAGGCTGAGGCTGAGGCTGAGGCT 3240
DB 3181 CCAAGGGCTGAGCTGAGGCTGAGGCTGAGGCTGAGGCTGAGGCTGAGGCTGAGGCTGAGGCT 3240
QY 3241 CTGTGGAAGAGACAGATTTGTTGAGGCTGAGGCTGAGGCTGAGGCTGAGGCTGAGGCT 3300
DB 3241 CTGTGGAAGAGACAGATTTGTTGAGGCTGAGGCTGAGGCTGAGGCTGAGGCTGAGGCT 3300
QY 3301 AGTGCATTAATGAGGCTGAGGCTGAGGCTGAGGCTGAGGCTGAGGCTGAGGCTGAGGCT 3360
DB 3301 AGTGCATTAATGAGGCTGAGGCTGAGGCTGAGGCTGAGGCTGAGGCTGAGGCTGAGGCT 3360
QY 3361 GGGCATCTGTCAGGAGGCTGAGGCTGAGGCTGAGGCTGAGGCTGAGGCTGAGGCTGAGGCT 3420
DB 3361 GGGCATCTGTCAGGAGGCTGAGGCTGAGGCTGAGGCTGAGGCTGAGGCTGAGGCTGAGGCT 3420
QY 3421 TGGAGACAAGCGGGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 3480
DB 3421 TGGAGACAAGCGGGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 3480
QY 3481 CATACACCACTTCAATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 3540
DB 3481 CATACACCACTTCAATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 3540
QY 3541 AACCCAGCTCTGAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 3600
DB 3541 AACCCAGCTCTGAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 3600
QY 3601 GCTCATATTTTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 3660
DB 3601 GCTCATATTTTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 3660
QY 3661 TGTGTCATTAACCCAGATATTAACAATCTTGGACAGGCAATATATATCTTATATCA 3720
DB 3661 TGTGTCATTAACCCAGATATTAACAATCTTGGACAGGCAATATATATCTTATATCA 3720
QY 3721 CTGTGTCATTAACCCAGATATTAACAATCTTGGACAGGCAATATATATCTTATATCA 3780
DB 3721 CTGTGTCATTAACCCAGATATTAACAATCTTGGACAGGCAATATATATCTTATATCA 3780
QY 3781 GATGTCACACATCAAGAGCTGCTGAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 3840
DB 3781 GATGTCACACATCAAGAGCTGCTGAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 3840
QY 3841 CCAGGCTGAGAGAGGCTGAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 3900
DB 3841 CCAGGCTGAGAGAGGCTGAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 3900
```

DB 3841 CCAGGCTGGAGCAACCGCTAGTGAAGTGGCCATATGACCTGTTAAATTTTTTAAAT 3900
QY 3901 ATTGTGTTAAACATGCGATTTATCAAAAGTTAAAGAGAGCACTTACTGAAAAACT 3960
DB 3901 ATTGTGTTAAACATGCGATTTATCAAAAGTTAAAGAGAGCACTTACTGAAAAACT 3960
QY 3961 ATGTGAACTACCTGTTTAAACATTTCTGCTGCAAGTCAAGATTCACCAAGTTTCA 4020
DB 3961 ATGTGAACTACCTGTTTAAACATTTCTGCTGCAAGTCAAGATTCACCAAGTTTCA 4020
QY 4021 AGCTTCAGATTTTATATTAAGAGCAAAAGCAAAATATATGATATTAATATC 4080
DB 4021 AGCTTCAGATTTTATATTAAGAGCAAAAGCAAAATATATGATATTAATATC 4080
QY 4081 TGGTGTATTTGATTTTAAATTAAGATTAATCAAGATTAATTTGTTAAATATG 4140
DB 4081 TGGTGTATTTGATTTTAAATTAAGATTAATCAAGATTAATTTGTTAAATATG 4140
QY 4141 TATATTTTGTATATTTTATTTTGTACTTACTGCTTGTGAAAGATTAAGAAGTG 4200
DB 4141 TATATTTTGTATATTTTATTTTGTACTTACTGCTTGTGAAAGATTAAGAAGTG 4200
QY 4201 GTAAGAAAGTACTGAATGTTGATTAAGTCAAGTCAATATTAAGAACTTTTATATC 4260
DB 4201 GTAAGAAAGTACTGAATGTTGATTAAGTCAAGTCAATATTAAGAACTTTTATATC 4260
QY 4261 AAAAAAAAAAAAAAAAAA 4279
DB 4261 AAAAAAAAAAAAAAAAAA 4279

RESULT 2
AADO3504
ID AADO3504 standard; cDNA; 4279 BP.
XX
AC AADO3504;
XX
DT 13-JUN-2001 (first entry)
XX
DE Dog P-glycoprotein (Pgp) allelic variant (Genotype A) cDNA.
XX
KW Dog: P-glycoprotein allelic variant; multidrug transporter; MDRI.
XX
KW drug bioavailability; transgenic animal; genetic model; ss.
XX
OS Canis familiaris.
XX
FH Key location/Qualifiers
FT CDS 17..3862
FT FT /*tag= a
FT FT /product= "Dog P-glycoprotein (Pgp) allelic variant
FT FT (Genotype A) protein"
FT FT replace (607, C)
FT FT /*tag= b
XX
PN WO200123540-A2.
XX
PD 05-APR-2001.
XX
PE 28-SEP-2000; 2000MO-US26767.
XX
PR 28-SEP-1999; 99US-0156510.
XX
PA (GENT-) GENEST CORP.
XX
PI Stocker PJ, Steimel-Crespi DR, Crespi CL, Reif NC, Patten CJ;
XX
DR WPI: 2001-235373/24.
XX
DR P-PSDB: AAED0308.
XX
PT New dog P-glycoproteins (Pgp) and their encoding nucleic acids, useful
PT for determining the bioavailability of drugs and for screening for dog
PT Pgp inhibitors -
XX

PS Claim 9; Page 85-90; 11pp; English.

XX The invention relates to dog P-glycoprotein (Pgp) also referred
CC as multidrug transporter (MDRI) and nucleic acids encoding them.
CC The invention also includes fragments and biologically functional
CC variants of dog P-glycoprotein. Pgp and their nucleic acids are
CC useful for determining the bioavailability of drugs and for
CC screening Pgp inhibitors. They are useful for the diagnosis and
CC treatment of conditions characterised by Pgp activity, by
CC reducing or increasing Pgp activity in a cell. Pgp nucleic acids
CC are used as oligonucleotide probes. Complements of Pgp nucleic
CC acids are useful as antisense oligonucleotides, to induce a Pgp
CC "knockout" phenotype. They are used to prepare a non-human
CC transgenic animal, which are valuable as genetic models for
CC human diseases.
CC The present sequence is dog P-glycoprotein (Pgp) allelic variant
CC (Genotype A) cDNA. The Pgp enzyme functions as an efflux pump
CC exporting small molecules across the cell membrane. This enzyme
CC is a member of the ABC transporter family.

SO Sequence 4279 BP; 1295 A; 833 C; 1008 G; 1143 T; 0 other;

Query Match 100.0%; Score 4277.4; DB 22; Length 4279;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 4278; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 GGAGCGGAGGTGCGATGATCTGGAAGAGCCGTAAGGGAGTGCAGAGAGAACTT 60
DB 1 GGAGCGGAGGTGCGATGATCTGGAAGAGCCGTAAGGGAGTGCAGAGAGAACTT 60
QY 61 CTGGAATAATGGGCAAAAAAAAAATGAAAGAGAGAGAGAGAGAGAGAGAGAGAGAG 120
DB 61 CTGGAATAATGGGCAAAAAAAAAATGAAAGAGAGAGAGAGAGAGAGAGAGAGAGAG 120
QY 121 CACGTTTGCATTTGTTGCTATTCATTAATTTGCTGATAGTGTATATGTTGGTGGGAC 180
DB 121 CACGTTTGCATTTGTTGCTATTCATTAATTTGCTGATAGTGTATATGTTGGTGGGAC 180
QY 181 AATGGCTGCGATCATCGATGAGAGCTGACCTCTCATGATGCTGATTTTGGAAACAT 240
DB 181 AATGGCTGCGATCATCGATGAGAGCTGACCTCTCATGATGCTGATTTTGGAAACAT 240
QY 241 GACGATAGCTTTGCAAAATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 300
DB 241 GACGATAGCTTTGCAAAATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 300
QY 241 GACGATAGCTTTGCAAAATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 300
DB 241 GACGATAGCTTTGCAAAATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 300
QY 301 TGAAGATTTACAGACATATACACATTTTATCATCAACATCTGAGAGAGAGAGAGAG 360
DB 301 TGAAGATTTACAGACATATACACATTTTATCATCAACATCTGAGAGAGAGAGAGAG 360
QY 361 GTATGCTTATTTTATTAAGTGGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 420
DB 361 GTATGCTTATTTTATTAAGTGGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 420
QY 421 TTTCATTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 480
DB 421 TTTCATTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 480
QY 481 TGTATATATGAG 540
DB 481 TGTATATATGAG 540
QY 541 CCGGCTCAG 600
DB 541 CCGGCTCAG 600
QY 601 CTTTCATCTCAATATGCAACATTTTACCGGTTTATATGTTGGGTTTACAGTGGTTGAA 660
DB 601 CTTTCATCTCAATATGCAACATTTTACCGGTTTATATGTTGGGTTTACAGTGGTTGAA 660
QY 661 GCTAACCTTTGATTTTGGCCATAGCCCTGTTCTTGGACCTTTCAGCCGCTATGGGC 720
DB 661 GCTAACCTTTGATTTTGGCCATAGCCCTGTTCTTGGACCTTTCAGCCGCTATGGGC 720

QY 721 AAGATACATCTTCAATTAAGAACTGTGGCCATGCAAAAGCTGAGCAGT 780
DB 721 AAGATACATCTTCAATTAAGAACTGTGGCCATGCAAAAGCTGAGCAGT 780
QY 781 AGCTAAGAAAGTCTTAGCAGCAATCAGAACTGATGCTGGAGAGCAAAAGAAAG 840
DB 781 AGCTAAGAAAGTCTTAGCAGCAATCAGAACTGATGCTGGAGAGCAAAAGAAAG 840
QY 841 ACTTGAAGAGTACAAACAAAAATTTAGAAAGCTAAAGAAATTTGGATTAAGAAAGCTAT 900
DB 841 ACTTGAAGAGTACAAACAAAAATTTAGAAAGCTAAAGAAATTTGGATTAAGAAAGCTAT 900
QY 901 CACGGCCAAACATTTCTATTTGTCGCCCTTCTTATTTGATCTATGATCTATGCTGGC 960
DB 901 CACGGCCAAACATTTCTATTTGTCGCCCTTCTTATTTGATCTATGATCTATGCTGGC 960
QY 961 TTTCTGATGAGGACCTTCTGCTCTGCACTGATATTTCTATTTGCAAGTACTGAC 1020
DB 961 TTTCTGATGAGGACCTTCTGCTCTGCACTGATATTTCTATTTGCAAGTACTGAC 1020
QY 1021 TGTCTTTTCTGTATTAATTTGGGCTTTTGTATTTGAGAGGATCCCAAGCTTGA 1080
DB 1021 TGTCTTTTCTGTATTAATTTGGGCTTTTGTATTTGAGAGGATCCCAAGCTTGA 1080
QY 1081 AGCATTTGCAAGCAGAGAGGAGCAGCTTATGAAATCTTCAAGATATTTGCAATTAAC 1140
DB 1081 AGCATTTGCAAGCAGAGAGGAGCAGCTTATGAAATCTTCAAGATATTTGCAATTAAC 1140
QY 1141 AAGCATTTGCAAGCAGAGAGGAGCAGCTTATGAAATCTTCAAGATATTTGCAATTAAC 1200
DB 1141 AAGCATTTGCAAGCAGAGAGGAGCAGCTTATGAAATCTTCAAGATATTTGCAATTAAC 1200
QY 1201 ATTCAAAATTTGCTCACTTCACTTCTGCAAAAGAAATTTAAGATTTTAAAGGCTCT 1260
DB 1201 ATTCAAAATTTGCTCACTTCACTTCTGCAAAAGAAATTTAAGATTTTAAAGGCTCT 1260
QY 1261 CAACCTGAAGTTTCAGAGTGGGCAAGCAGTGGGCTGGTGGGAAAGCTGGCGGAA 1320
DB 1261 CAACCTGAAGTTTCAGAGTGGGCAAGCAGTGGGCTGGTGGGAAAGCTGGCGGAA 1320
QY 1321 GAGCAGAGCCTGAGTGGAGAGGCTCTATGACCCCAAGATGGCTGTGTAT 1380
DB 1321 GAGCAGAGCCTGAGTGGAGAGGCTCTATGACCCCAAGATGGCTGTGTAT 1380
QY 1381 TGATGACAGAGACATTTAGAACATTAATGTAAAGCATCTTGGGAAATTTACTGTGTGT 1440
DB 1381 TGATGACAGAGACATTTAGAACATTAATGTAAAGCATCTTGGGAAATTTACTGTGTGT 1440
QY 1441 GAGTCAGAGAGCTGT 1500
DB 1441 GAGTCAGAGAGCTGT 1500
QY 1501 AAATGTACCATGATGATGATTTGAAAGCTTTTAAAGAAAGCAATGCTATGATTTTAT 1560
DB 1501 AAATGTACCATGATGATGATTTGAAAGCTTTTAAAGAAAGCAATGCTATGATTTTAT 1560
QY 1561 CATGAATCTACCTTAATTAATTTGACACTGTGTGTGAGAGAGAGGCGCCACCTAGTGG 1620
DB 1561 CATGAATCTACCTTAATTAATTTGACACTGTGTGTGAGAGAGAGGCGCCACCTAGTGG 1620
QY 1621 TGGACAGAAACAGAGATGAGCTTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1680
DB 1621 TGGACAGAAACAGAGATGAGCTTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1680
QY 1681 GCTGGATGAGGCAAGCTGAGCTGTGAGACATGAAAGTGAAGCAGAGTGTGAGTGGCCCT 1740
DB 1681 GCTGGATGAGGCAAGCTGAGCTGTGAGACATGAAAGTGAAGCAGAGTGTGAGTGGCCCT 1740
QY 1741 GGATTAAGGCAAGAAAGGCGGAGCTAGCATTTGATCTATGCTTGTGTGTGTGTGTGTGT 1800
DB 1741 GGATTAAGGCAAGAAAGGCGGAGCTAGCATTTGATCTATGCTTGTGTGTGTGTGTGTGT 1800

QY 1801 TAATGCCATGTCATCTGCTGTTTGTATGATGAGTCAATTTGTGGAGAAAGAAATCATGA 1860
DB 1801 TAATGCCATGTCATCTGCTGTTTGTATGATGAGTCAATTTGTGGAGAAAGAAATCATGA 1860
QY 1861 TGAATCTAGAAAGAGAGGCTTACTTCAAACTTGTACAAATGACAGCAGAGAGAA 1920
DB 1861 TGAATCTAGAAAGAGAGGCTTACTTCAAACTTGTACAAATGACAGCAGAGAGAA 1920
QY 1921 TGAATTTGAGTTACAAATTTGCACTGTGTGATCCAAAGTGAAGTGTATGCTTGAAT 1980
DB 1921 TGAATTTGAGTTACAAATTTGCACTGTGTGATCCAAAGTGAAGTGTATGCTTGAAT 1980
QY 1981 GTCTCCAAAGATTCAGAGTCCAGTTTAAATTAATAAAGAAATCACTGCGAGAGATPACA 2040
DB 1981 GTCTCCAAAGATTCAGAGTCCAGTTTAAATTAATAAAGAAATCACTGCGAGAGATPACA 2040
QY 2041 TGCACCAAGAGCCAGACAGAAAGCTTGTACAAAAGAGAGCTTGAATGAGATGTACC 2100
DB 2041 TGCACCAAGAGCCAGACAGAAAGCTTGTACAAAAGAGAGCTTGAATGAGATGTACC 2100
QY 2101 TCCAGTTTCTTGTGAGAGTCTGAACTGAACTCACTGATGCTTATTTGTGTGT 2160
DB 2101 TCCAGTTTCTTGTGAGAGTCTGAACTGAACTCACTGATGCTTATTTGTGTGT 2160
QY 2161 TGTATTAATTTTGTCTATTTAATTAAGAGAGGCTGCAACAGATTTTCAATTAATTTTC 2220
DB 2161 TGTATTAATTTTGTCTATTTAATTAAGAGAGGCTGCAACAGATTTTCAATTAATTTTC 2220
QY 2221 AAGGATTAAGAGATCTTACCCGAGATGAGATCTGTAACAAAGCAGAAATAGTAA 2280
DB 2221 AAGGATTAAGAGATCTTACCCGAGATGAGATCTGTAACAAAGCAGAAATAGTAA 2280
QY 2281 CATGTTTCTGTATTTGTCTAGTCTTGAATATTTCTTTATTTACATTTTCCCTCA 2340
DB 2281 CATGTTTCTGTATTTGTCTAGTCTTGAATATTTCTTTATTTACATTTTCCCTCA 2340
QY 2341 GGGCTTCAATTTGGCAAGCTGGGAGATCTCTCACTAAGCGGCTTGATATCATGTTT 2400
DB 2341 GGGCTTCAATTTGGCAAGCTGGGAGATCTCTCACTAAGCGGCTTGATATCATGTTT 2400
QY 2401 CAGATCATGCTGAGACAGATGTCAGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 2460
DB 2401 CAGATCATGCTGAGACAGATGTCAGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 2460
QY 2461 ATTGCAACCAAGGCTTGCATATGATGCTGCAATTTAAAGGGCTATAGTTTCCAGGCT 2520
DB 2461 ATTGCAACCAAGGCTTGCATATGATGCTGCAATTTAAAGGGCTATAGTTTCCAGGCT 2520
QY 2521 TGTGTCTATTAACCAAGATATGCAAAATCTTGGACAGAGCATTAATATTTCTTATCTA 2580
DB 2521 TGTGTCTATTAACCAAGATATGCAAAATCTTGGACAGAGCATTAATATTTCTTATCTA 2580
QY 2581 TGGTTGGCATTTAAACATTTTACTCTTGAATTTGACCATCATTTGCAATTTGAGAGAGT 2640
DB 2581 TGGTTGGCATTTAAACATTTTACTCTTGAATTTGACCATCATTTGCAATTTGAGAGAGT 2640
QY 2641 TGTGAAATGAAATTTTCTGTGACACAGCAGTGAAGTAAAGAGAGCTTGAAGAGC 2700
DB 2641 TGTGAAATGAAATTTTCTGTGACACAGCAGTGAAGTAAAGAGAGCTTGAAGAGC 2700
QY 2701 TGGGAGAGATTTGACAGAAAGCAATGCAAAATCTTCCAACTGTTTCTTGTGACTGGGA 2760
DB 2701 TGGGAGAGATTTGACAGAAAGCAATGCAAAATCTTCCAACTGTTTCTTGTGACTGGGA 2760
QY 2761 GCAGAGATTTGAATCATGATGACAGAGTTTGAAGTACCATTAACAGAACTTTTGGAG 2820
DB 2761 GCAGAGATTTGAATCATGATGACAGAGTTTGAAGTACCATTAACAGAACTTTTGGAG 2820
QY 2821 GAAAGACACATCTTGGGCTGTCAATTTGTATACCCAGGCAATGATGATTTTTCCTA 2880
DB 2821 GAAAGACACATCTTGGGCTGTCAATTTGTATACCCAGGCAATGATGATTTTTCCTA 2880
QY 2881 TGTGTGCTGTTTCCGTTTGT 2940

```

|||||
Db 2881 TGTGGCTTTTCGGTTTGGGCTACTGTGGCAAAATGATTCATGAACTTCAGGA 2940
|||
QY 2941 TGTCTTTTGTATTCACGATATGTCTTTGGGCCATGGCAGTGGGCGAGCTCAATTC 3000
|||||
Db 2941 TGTCTTTTGTATTCACGATATGTCTTTGGGCCATGGCAGTGGGCGAGCTCAATTC 3000
|||
QY 3001 ATTTGCTCCGACTATGCCAAAGATATCAGACCCCACTCATCATGATATTTGA 3060
|||||
Db 3001 ATTTGCTCCGACTATGCCAAAGATATCAGACCCCACTCATCATGATATTTGA 3060
|||
QY 3061 AAAAGCCCTGTGATGACAGCTACAGCCCTCAGCGCCCTCAAGCCCAATAGCTTGAAG 3120
|||||
Db 3061 AAAAGCCCTGTGATGACAGCTACAGCCCTCAGCGCCCTCAAGCCCAATAGCTTGAAG 3120
|||
QY 3121 AAATGACATTTAATGAGGTCGTGTTCAACTATCCCACTGACACAGACATCCCGTCT 3180
|||||
Db 3121 AAATGACATTTAATGAGGTCGTGTTCAACTATCCCACTGACACAGACATCCCGTCT 3180
|||
QY 3181 CCAGGGGCTGAGCCTGAGGTGAAGAGGCCGAGAGCGCGGCTGTAGGTAGCAGTGG 3240
|||||
Db 3181 CCAGGGGCTGAGCCTGAGGTGAAGAGGCCGAGAGCGCGGCTGTAGGTAGCAGTGG 3240
|||
QY 3241 CTGTGGGAAGGACAGTTGTGATGAGTCTGAGCGCTTCTATGACCCCTTGCTGCTTC 3300
|||||
Db 3241 CTGTGGGAAGGACAGTTGTGATGAGTCTGAGCGCTTCTATGACCCCTTGCTGCTTC 3300
|||
QY 3301 AGTGTAAATGATGGCAAGATAAAGCACCTGAATGTCATGAGTCCGAGACACCT 3360
|||||
Db 3301 AGTGTAAATGATGGCAAGATAAAGCACCTGAATGTCATGAGTCCGAGACACCT 3360
|||
QY 3361 GGGCATGCTGTCTGAGAGCCATCTGTTTGAATGACGATTTGCCGAAGACATTTGCC 3420
|||||
Db 3361 GGGCATGCTGTCTGAGAGCCATCTGTTTGAATGACGATTTGCCGAAGACATTTGCC 3420
|||
QY 3421 TGGAGACAACAGCCGCTGATCATCATGAGAGATTAATGAGGTCACCAAGAGGCGCA 3480
|||||
Db 3421 TGGAGACAACAGCCGCTGATCATCATGAGAGATTAATGAGGTCACCAAGAGGCGCA 3480
|||
QY 3481 CATACACCACTTCATGAGACACTCCCTGAGAAATACAAACAGAGTAGAGACAAAG 3540
|||||
Db 3481 CATACACCACTTCATGAGACACTCCCTGAGAAATACAAACAGAGTAGAGACAAAG 3540
|||
QY 3541 AACCCAGCTCTGTGTGCGCAGAAACAGCGCATTTGCCATAGCTGGCGCTTTAGACA 3600
|||||
Db 3541 AACCCAGCTCTGTGTGCGCAGAAACAGCGCATTTGCCATAGCTGGCGCTTTAGACA 3600
|||
QY 3601 GCCCTATTTTGTGATGAGCTACATCAGCTCGATACAGAAAGTGAAGAGT 3660
|||||
Db 3601 GCCCTATTTTGTGATGAGCTACATCAGCTCGATACAGAAAGTGAAGAGT 3660
|||
QY 3661 TGTCCAGAAGCCCTGGACAAAGCCAGAGAGCGCGACCTGATGATGCGCCACCG 3720
|||||
Db 3661 TGTCCAGAAGCCCTGGACAAAGCCAGAGAGCGCGACCTGATGATGCGCCACCG 3720
|||
QY 3721 CTGTCCACCATCCAGATCCAGATTTAATAGTGTTCAGAAATGCGAAAGTCAAGGA 3780
|||||
Db 3721 CTGTCCACCATCCAGATCCAGATTTAATAGTGTTCAGAAATGCGAAAGTCAAGGA 3780
|||
QY 3781 GCATGCGACATCAGATCAGTGTGCGCCAGAAAGCATCTATTTTCCATGGTCAAGT 3840
|||||
Db 3781 GCATGCGACATCAGATCAGTGTGCGCCAGAAAGCATCTATTTTCCATGGTCAAGT 3840
|||
QY 3841 CCAGGCTGAGACAAAGCGCTAGTGAATGCGCATATGAGCTGTAAATATTTTAAT 3900
|||||
Db 3841 CCAGGCTGAGACAAAGCGCTAGTGAATGCGCATATGAGCTGTAAATATTTTAAT 3900
|||
QY 3901 ATTTGTGTTAAACATGATTAATCAAGTTAAAGGTGACACTTACTGGAAGAACT 3960
|||||
Db 3901 ATTTGTGTTAAACATGATTAATCAAGTTAAAGGTGACACTTACTGGAAGAACT 3960
|||
QY 3961 ATGTAGAACTACCTGTTTAACATTTCTGTGCAACTGAAGATCAATCCACCAAGTTCAG 4020
|||||

```

```

Db 3961 ATGTAGAACTACCTGTTTAACATTTCTGCTGCAACTGAGATCATTTCCACCAAGTTCAG 4020
QY 4021 AGCTTCAGATTTTAAATTAATTAAGAACCAAGAACATTAATCTGATGATTAATTAATAG 4080
Db 4021 AGCTTCAGATTTTAAATTAATTAAGAACCAAGAACATTAATCTGATGATTAATTAATAG 4080
QY 4081 TGTGTGTTAATTCATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 4140
Db 4081 TGTGTGTTAATTCATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 4140
QY 4141 TATTAATTTTGTGTTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 4200
Db 4141 TATTAATTTTGTGTTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 4200
QY 4201 GTAAAAAGTACGATGTTGAAATTAAGTACGCTATTAATTAATTAATTAATTAATTC 4260
Db 4201 GTAAAAAGTACGATGTTGAAATTAAGTACGCTATTAATTAATTAATTAATTAATTC 4260
QY 4261 AAAAAAAAAAAAAAAAAAAAAA 4279
Db 4261 AAAAAAAAAAAAAAAAAAAAAA 4279

```

RESULT 3
AAD03505
ID AAD03505 standard; cDNA; 4279 BP.

AC AAD03505;

DT 13-JUN-2001 (first entry)

DE Dog P-glycoprotein (PGP) allelic variant (Genotype B) cDNA.

KW Dog; P-glycoprotein allelic variant; PGP; multidrug transporter; MDRL; drug bioavailability; transgenic animal; genetic model; ss.

OS Canis familiaris.

XX Key Location/Qualifiers

XX CDS 17..3862

XX /tag= a
/product= "Dog P-glycoprotein (PGP) allelic variant
(Genotype B) protein"

XX allele
replace (91, T)

XX /tag= b
replace (607, C)

XX allele
/tag= c

XX WO200123540-A2.

XX 05-APR-2001.

XX 28-SEP-2000; 2000WO-US26767.

XX 28-SEP-1999; 99US-0156510.

XX (GENT-) GENTEST CORP.

XX Stocker PJ, Steinel-crespi DR, Crespi CL, Reif TC, Patten CJ;

XX WPI; 2001-235373/24.

XX P-PSDB; AAE00309.

XX New dog P-glycoproteins (PGP) and their encoding nucleic acids, useful
for determining the bioavailability of drugs and for screening for dog

XX PGP inhibitors -

XX Claim 9; Page 93-99; 11pp; English.

XX The invention relates to dog P-glycoprotein (PGP) also referred
as multidrug transporter (MDRL) and nucleic acids encoding them.
The invention also includes fragments and biologically functional
variants of dog P-glycoprotein. PGP and their nucleic acids are

CC useful for determining the bioavailability of drugs and for
CC screening PGP inhibitors. They are useful for the diagnosis and
CC treatment of conditions characterised by PGP activity, by
CC reducing or increasing PGP activity in a cell. PGP nucleic acids
CC are used as oligonucleotide probes. Complements of PGP nucleic
CC 'knockout' phenotype. They are used to prepare a non-human
CC transgenic animal, which are valuable as genetic models for
CC human diseases.
CC The present sequence is dog P-glycoprotein (PGP) allelic variant
CC (genotype B) cDNA. The PGP enzyme functions as an efflux pump
CC exporting small molecules across the cell membrane. This enzyme
CC is a member of the ABC transporter family.
CC
xx

Sequence 4279 BP; 1296 A; 833 C; 1008 G; 1142 T; 0 other;

Query Match 99.9%; Score 4275.8; DB 22; Length 4279;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 4277; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

```

OY 1 GGAGCCGGAGTGGGATGGATCCTGAAGAGCCGTGAAGGGAGTGCAGAGAAGACTT 60
Db 1 GGAGCCGGAGTGGGATGGATCCTGAAGAGCCGTGAAGGGAGTGCAGAGAAGACTT 60
OY 61 CTGAAAAATGGCAAAAAAGTAAAAAATGAGAAAGAAAAAGAAACCAACTGTCAG 120
Db 61 CTGAAAAATGGCAAAAAAGTAAAAAATGAGAAAGAAAAAGAAACCAACTGTCAG 120
OY 121 CAGCTTGCAGATGTTGCTATTCAAATTCGCTATGATGATGATGATGATGATGATG 180
Db 121 CAGCTTGCAGATGTTGCTATTCAAATTCGCTATGATGATGATGATGATGATGATG 180
OY 181 AATGGCTGCATCATCCATGAGAGTGCACCTCCCTCATGATGCTGTTTGGAAACAT 240
Db 181 AATGGCTGCATCATCCATGAGAGTGCACCTCCCTCATGATGCTGTTTGGAAACAT 240
OY 241 GACAGATAGCTTTCGCAAAATGCAAGAAATTCAGAAACAACTTCCAGTAAATTA 300
Db 241 GACAGATAGCTTTCGCAAAATGCAAGAAATTCAGAAACAACTTCCAGTAAATTA 300
OY 301 TGAAGATTTAGCAACATATACAAACATTTATACACATCTGGAGAGGAAATGACAC 360
Db 301 TGAAGATTTAGCAACATATACAAACATTTATACACATCTGGAGAGGAAATGACAC 360
OY 361 GATGCTTATTTACAGTGGATCGTGGCTGCTGGCTGCTGGCTGCTGGCTGCTGGCT 420
Db 361 GATGCTTATTTACAGTGGATCGTGGCTGCTGGCTGCTGGCTGCTGGCTGCTGGCT 420
OY 421 TTCAATTCGGTGGCTGGCAGCAGAGACAGATACCTCAAAATTTAGAAACAAATTTTCA 480
Db 421 TTCAATTCGGTGGCTGGCAGCAGAGACAGATACCTCAAAATTTAGAAACAAATTTTCA 480
OY 481 TGCTATCATCGCAGAGAGATGGCTGGTTTGACGTGCATGACGTTGGGAGCCTTAAC 540
Db 481 TGCTATCATCGCAGAGAGATGGCTGGTTTGACGTGCATGACGTTGGGAGCCTTAAC 540
OY 541 CCGGCTACAGAGAGATGTCGCAAAATCAATGAAGAAATTTGGCAGAAATTTGGATTT 600
Db 541 CCGGCTACAGAGAGATGTCGCAAAATCAATGAAGAAATTTGGCAGAAATTTGGATTT 600
OY 601 CTTTCACCTAATAGCAACATTTTTCACCGGTTTATAGGGGTTTACACGTGGTGA 660
Db 601 CTTTCACCTAATAGCAACATTTTTCACCGGTTTATAGGGGTTTACACGTGGTGA 660
OY 661 GCTAACCCCTTGATTTGGCAGTACGCCCTGTTCTTGAGCTTTCACCGGCTGAGG 720
Db 661 GCTAACCCCTTGATTTGGCAGTACGCCCTGTTCTTGAGCTTTCACCGGCTGAGG 720
OY 721 AAAGATATATCTTATTTAGATGAAGAACTGTTGGCTATGCAAAAGCTGGAGAGT 780
Db 721 AAAGATATATCTTATTTAGATGAAGAACTGTTGGCTATGCAAAAGCTGGAGAGT 780
OY 781 AACTGAAGAGCTTTAGAGAGCAATCAGAACTGTGATGCTTTGGAGAGCAAAAGAGA 840
```

```

Db 781 AACTGAAGAGCTTTAGAGAGCAATCAGAACTGTGATGCTTTGGAGAGCAAAAGAGA 840
OY 841 ACTTGAAGAGTACAAACAAATTTAGAGAGCTTAAAGAAATTTGGATTAAGAAAGCTAT 900
Db 841 ACTTGAAGAGTACAAACAAATTTAGAGAGCTTAAAGAAATTTGGATTAAGAAAGCTAT 900
OY 901 CAGGCGCAACATTTCTATTTGCTGGCTTTCTTATGATCTATGATCAATATGCTGTC 960
Db 901 CAGGCGCAACATTTCTATTTGCTGGCTTTCTTATGATCTATGATCAATATGCTGTC 960
OY 961 TTTCTGATGAGAGCTGCTGCTGGCTCTCCAGTGAATATTCATTTGACAAAGTACTAC 1020
Db 961 TTTCTGATGAGAGCTGCTGCTGGCTCTCCAGTGAATATTCATTTGACAAAGTACTAC 1020
OY 1021 TGTCTTTCTTTCTGATTAATTTGGGCTTTTATGATTTGAGAGAGCATCCCAAGCAT 1080
Db 1021 TGTCTTTCTTTCTGATTAATTTGGGCTTTTATGATTTGAGAGAGCATCCCAAGCAT 1080
OY 1081 AGCATTTGCAAGCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1140
Db 1081 AGCATTTGCAAGCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1140
OY 1141 AAGCATTTGACAGCTATTCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1200
Db 1141 AAGCATTTGACAGCTATTCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1200
OY 1201 ATTCAAAATATGTTCACTTACCTGATTTACCTTTCGAAAGAAAGTTAAGATCTTAAAG 1260
Db 1201 ATTCAAAATATGTTCACTTACCTGATTTACCTTTCGAAAGAAAGTTAAGATCTTAAAG 1260
OY 1261 CAACCTGAAAGGTTCAAGAGTGGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1320
Db 1261 CAACCTGAAAGGTTCAAGAGTGGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1320
OY 1321 GAGCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1380
Db 1321 GAGCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1380
OY 1381 TGATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1440
Db 1381 TGATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1440
OY 1441 GAGTACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1500
Db 1441 GAGTACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1500
OY 1501 AAATGTACCATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1560
Db 1501 AAATGTACCATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1560
OY 1561 CATGAATCTACCTAATTAATTTTACACTCTGTTTGGAGAGAGAGAGAGAGAGAGAGAG 1620
Db 1561 CATGAATCTACCTAATTAATTTTACACTCTGTTTGGAGAGAGAGAGAGAGAGAGAGAG 1620
OY 1621 TGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1680
Db 1621 TGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1680
OY 1681 GCTGGATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1740
Db 1681 GCTGGATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1740
OY 1741 GATTAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1800
Db 1741 GATTAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1800
OY 1801 TAAATCCGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1860
Db 1801 TAAATCCGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1860
OY 1861 TGAATCTAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1920
Db 1861 TGAATCTAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1920
```

Db	1861	TGAACTCATGAAGAAGAGGGCATTTACTTCTCAAACTGTCTCAATATGACAGACAGGAAA	1920
Qy	1921	TGAATTTGAGTTGAAAAATGCCCTGGTGAATCCAAAATGCAAAATGATGGCTTGGAAAT	1980
Db	1921	TGAATTTGAGTTGAAAAATGCCCTGGTGAATCCAAAATGCAAAATGATGGCTTGGAAAT	1980
Qy	1981	GTCCCAAAAGATTCCAGGGTCCAGTTTAAATAAAAGACATCAACTCGCAGAGATATCA	2040
Db	1981	GTCTCCAAAAGATTCCAGGGTCCAGTTTAAATAAAAGACATCAACTCGCAGAGATATCA	2040
Qy	2041	TGCACCAACAGGCCAAGACAGAAAGCTGGTACAAAAGAGGACCTTGAATGAAATGTAC	2100
Db	2041	TGCACCAACAGGCCAAGACAGAAAGCTGGTACAAAAGAGGACCTTGAATGAAATGTAC	2100
Qy	2101	TCCAGTTTCCTTCTGGAGAGATTCTGAAAGCTGAACCTCAACTGAATGGCTTATTTTGGT	2160
Db	2101	TCCAGTTTCCTTCTGGAGAGATTCTGAAAGCTGAACCTCAACTGAATGGCTTATTTTGGT	2160
Qy	2161	TGGTATATTTTGTGCTATTTATAACGGAGGCTGCAACAGACATTTTCATATATTTTTC	2220
Db	2161	TGGTATATTTTGTGCTATTTATAACGGAGGCTGCAACAGACATTTTCATATATTTTTC	2220
Qy	2221	AAGGATTAAGGACTTTTACCCGAGATGAGATCTGAAACAAACAGACAAATGTAA	2280
Db	2221	AAGGATTAAGGACTTTTACCCGAGATGAGATCTGAAACAAACAGACAAATGTAA	2280
Qy	2281	CATGTTTCTGTATTTGTTTCTAGTCCCTTGGAAATATTTCTTTATACATTTTTCCTCA	2340
Db	2281	CATGTTTCTGTATTTGTTTCTAGTCCCTTGGAAATATTTCTTTATACATTTTTCCTCA	2340
Qy	2341	GGGCTTCACATTTGGCAAGCTGGGGAGATCCCTCAAGCGGCTTGATACATGTTTT	2400
Db	2341	GGGCTTCACATTTGGCAAGCTGGGGAGATCCCTCAAGCGGCTTGATACATGTTTT	2400
Qy	2401	CAGATCCATCTGAGACAGAGATGTACGTGGTTGATGACCCCTAATAAACACACATGTGAGC	2460
Db	2401	CAGATCCATCTGAGACAGAGATGTACGTGGTTGATGACCCCTAATAAACACACATGTGAGC	2460
Qy	2461	ATTGACAAACAGGGTGGCCAAATGATCGGCTCAAGTTAAAGGGCTATAGTTCCAGGCT	2520
Db	2461	ATTGACAAACAGGGTGGCCAAATGATCGGCTCAAGTTAAAGGGCTATAGTTCCAGGCT	2520
Qy	2521	TGCTGTCAATACCCAGAAATATACCAATCTTGGGAAAGGCAATTTATATCTTTAATCTA	2580
Db	2521	TGCTGTCAATACCCAGAAATATACCAATCTTGGGAAAGGCAATTTATATCTTTAATCTA	2580
Qy	2581	TGGTTGGCAATTAACACTTTTACTCTTACCAATTTGATACCATTCATATGCAAGAGT	2640
Db	2581	TGGTTGGCAATTAACACTTTTACTCTTACCAATTTGATACCATTCATATGCAAGAGT	2640
Qy	2641	TGTTGAAATGAAATGTTGTCTGCAACAAGACATGAAAGATTAATAAGACTAGAGAGAC	2700
Db	2641	TGTTGAAATGAAATGTTGTCTGCAACAAGACATGAAAGATTAATAAGACTAGAGAGAC	2700
Qy	2701	TGGGAAAGATTGCTACAGAAAGCATCGAAAACTCCGAACGTGTTTCTTTGACTCGGGA	2760
Db	2701	TGGGAAAGATTGCTACAGAAAGCATCGAAAACTCCGAACGTGTTTCTTTGACTCGGGA	2760
Qy	2761	GCAGAGATTGAAATACATGATAGCAACAAGTTTGCAAGTACATACAGAAACCTTTGAG	2820
Db	2761	GCAGAGATTGAAATACATGATAGCAACAAGTTTGCAAGTACATACAGAAACCTTTGAG	2820
Qy	2821	GAAAGCAACATCTCGGGGCTCATTTTCTATACACCAGGCAATGATATTTTCTTA	2880
Db	2821	GAAAGCAACATCTCGGGGCTCATTTTCTATACACCAGGCAATGATATTTTCTTA	2880
Qy	2881	TGCTGGCTGTTTCCGGTTTGTTGTTGCTACTTGGTGGCAAAATGATCATGAACTTTCAG	2940
Db	2881	TGCTGGCTGTTTCCGGTTTGTTGTTGCTACTTGGTGGCAAAATGATCATGAACTTTCAG	2940
Qy	2941	TGTTCTTTTGGTATCTGCAGCTATTTGTTTGGTGCCTATGGCAATGGGGCAGGTCACTTC	3000
Db	2941	TGTTCTTTTGGTATCTGCAGCTATTTGTTTGGTGCCTATGGCAATGGGGCAGGTCACTTC	3000

QY	3001	ATTGTGCTCGCATATGATGCCAAAGCCAAAGTATCAGACGCCACGTCATCATGATCTTGA	3060
Db	3001	ATTGTGCTCGCATATGATGCCAAAGCCAAAGTATCAGACGCCACGTCATCATGATCTTGA	3060
QY	3061	AAAAAGCCCTCTGATTGACAGCTACACGCCCTCAGCGCCTCAAGCCAAATACGTTGGANAG	3120
Db	3061	AAAAAGCCCTCTGATTGACAGCTACACGCCCTCAGCGCCTCAAGCCAAATACGTTGGANAG	3120
QY	3121	AAATGTGCATTTATGAGAGTCGTGTTTCACTATATCCACTGCACACACATATCCCCCGGT	3180
Db	3121	AAATGTGCATTTATGAGAGTCGTGTTTCACTATATCCACTGCACACACATATCCCCCGGT	3180
QY	3181	CCAGGGGCTGAGCCTCGAGGTGAAAGAAAGGGCCAGACGCTGGCCCTCGTAGTACAGTGG	3240
Db	3181	CCAGGGGCTGAGCCTCGAGGTGAAAGAAAGGGCCAGACGCTGGCCCTCGTAGTACAGTGG	3240
QY	3241	CTGTGGGAAGAGCAGATGTTGTAGCTCTTAAGCGCTTCTTAACCCCTTGGTGGTTG	3300
Db	3241	CTGTGGGAAGAGCAGATGTTGTAGCTCTTAAGCGCTTCTTAACCCCTTGGTGGTTG	3300
QY	3301	AGTGCTAATATGAGCAAAAGATTAAGACACCTGAATGTCCAGTGGCTCCAGACACACT	3360
Db	3301	AGTGCTAATATGAGCAAAAGATTAAGACACCTGAATGTCCAGTGGCTCCAGACACACT	3360
QY	3361	GGGCATCGTGTCTCAGAGACCCATCTCTGTTTGCATCGACATTTGGCGGAACATTTGCTA	3420
Db	3361	GGGCATCGTGTCTCAGAGACCCATCTCTGTTTGCATCGACATTTGGCGGAACATTTGCTA	3420
QY	3421	TGAGAGCAACACCGGGTGCATGCATACATGAAGAAGATTTATGACGACGACCCAGGACCA	3480
Db	3421	TGAGAGCAACACCGGGTGCATGCATACATGAAGAAGATTTATGACGACGACCCAGGACCA	3480
QY	3481	CATACACCACTTCATCAGAGACATCCCTCGAATAATATACACCAAGATAGGAGACAAAG	3540
Db	3481	CATACACCACTTCATCAGAGACATCCCTCGAATAATATACACCAAGATAGGAGACAAAG	3540
QY	3541	AACCCACCTCTCGTGGGCCAGAAACAGCGCATTTGCCATAGCTCGGGCTCTTTTGAACA	3600
Db	3541	AACCCACCTCTCGTGGGCCAGAAACAGCGCATTTGCCATAGCTCGGGCTCTTTTGAACA	3600
QY	3601	GCCTCATATTTTCTTTTGGATGGAAGCTACATCAGCTCTGGATACAGAAAGTCAAAAGT	3660
Db	3601	GCCTCATATTTTCTTTTGGATGGAAGCTACATCAGCTCTGGATACAGAAAGTCAAAAGT	3660
QY	3661	TGTCCAAGAAAGCCCTGGACAAAGCCAGAGAAAGGCGCACCTGATTTGATGCCACACG	3720
Db	3661	TGTCCAAGAAAGCCCTGGACAAAGCCAGAGAAAGGCGCACCTGATTTGATGCCACACG	3720
QY	3721	CTGTGCCACATCCAGAAATCAGATTTAATAGTGTGTTTCAAGATGGCAAAAGTCAAAGA	3780
Db	3721	CTGTGCCACATCCAGAAATCAGATTTAATAGTGTGTTTCAAGATGGCAAAAGTCAAAGA	3780
QY	3781	GCATGGCACACATCACACAGCTGCTGGCCCGCAAGAAAGCATATTTTTCATGGGTCACTGT	3840
Db	3781	GCATGGCACACATCACACAGCTGCTGGCCCGCAAGAAAGCATATTTTTCATGGGTCACTGT	3840
QY	3841	CCAGGCTGAGCAAAAGCGTATGTAACCTGTGGCATATGAGCTGTTAAATTTTTTAAAT	3900
Db	3841	CCAGGCTGAGCAAAAGCGTATGTAACCTGTGGCATATGAGCTGTTAAATTTTTTAAAT	3900
QY	3901	ATTGTGTTAAACATGSCATTTATCAAAAGTTAAAGAGTGAAGCACTTACTGGAAAACT	3960
Db	3901	ATTGTGTTAAACATGSCATTTATCAAAAGTTAAAGAGTGAAGCACTTACTGGAAAACT	3960
QY	3961	ATGTACAACTACCTGTTTAAACATTTCTTGCTGCAACTGAAGATCAATCCACCAAGTTAG	4020
Db	3961	ATGTACAACTACCTGTTTAAACATTTCTTGCTGCAACTGAAGATCAATCCACCAAGTTAG	4020
QY	4021	AGTCCTCAGATTTTAAATTAAGGAACCAAAAGAAACATATCTGATGGAATAAATATAC	4080
Db	4021	AGTCCTCAGATTTTAAATTAAGGAACCAAAAGAAACATATCTGATGGAATAAATATAC	4080

RESULT 4	
AA003506	
ID	AA003506 standard; cDNA; 4279 BP.
XX	
AC	AA003506;
XX	
DT	13-JUN-2001 (first entry)
XX	
DE	Dog P-glycoprotein (PGP) allelic variant (Genotype D) cDNA.
XX	
KM	Dog; P-glycoprotein allelic variant; PGP; multidrug transporter;
XX	MDR1; drug bioavailability; transgenic animal; genetic model; ss.
OS	Canis familiaris.
XX	
FH	Key
FT	Location/Qualifiers
FT	17..3862
FT	/tag= a
FT	/product= "Dog P-glycoprotein (PGP) allelic variant
FT	(Genotype D) protein"
FT	replace (I, T)
FT	/tag= b
FT	replace (607, C)
FT	/tag= c
FT	replace (1001, T)
FT	/tag= c
FT	replace (3458, A)
FT	/tag= c
XX	
PN	WO200123540-A2.
XX	
PD	05-APR-2001.
XX	
PF	28-SEP-2000; 2000WO-US26767.
XX	
PR	28-SEP-1999; 99US-0156510.
XX	
PA	(GENT-) GENTEST CORP.
XX	
PI	Stocker PJ, Steinmetz-crespi DT, Crespi CL, Reif TC, Patten CJ;
XX	WPI: 2001-235373/24.
DR	P-Psdb; AAE00310.
XX	
XX	New dog P-glycoproteins (PGP) and their encoding nucleic acids, usef
PT	for determining the bioavailability of drugs and for screening for d
PT	PGP inhibitors -
XX	
PS	Claim 9; Page 102-107; 11pp; English.
XX	
CC	The invention relates to dog P-glycoprotein (PGP) also referred
CC	as multidrug transporter (MDR1) and nucleic acids encoding them.
CC	The invention also includes fragments and biologically functional
CC	variants of dog P-glycoprotein, PGP and their nucleic acids are
CC	useful for determining the bioavailability of drugs and for
CC	screening PGP inhibitors. They are useful for the diagnosis and

sequence 42/9 BP; 1296 A; 833 C; 1009 G; 1141 T; 0 other;

```
Best Local Similarity 99.98; Pred. No. 0; DB 22, Length 42/9,
Matches 4275; Conservative 0; w.t. 1.00
```

gaps; users

1
 2
 3
 4
 5
 6
 7
 8
 9
 10
 11
 12
 13
 14
 15
 16
 17
 18
 19
 20
 21
 22
 23
 24
 25
 26
 27
 28
 29
 30
 31
 32
 33
 34
 35
 36
 37
 38
 39
 40
 41
 42
 43
 44
 45
 46
 47
 48
 49
 50
 51
 52
 53
 54
 55
 56
 57
 58
 59
 60
 61
 62
 63
 64
 65
 66
 67
 68
 69
 70
 71
 72
 73
 74
 75
 76
 77
 78
 79
 80
 81
 82
 83
 84
 85
 86
 87
 88
 89
 90
 91
 92
 93
 94
 95
 96
 97
 98
 99
 100
 101
 102
 103
 104
 105
 106
 107
 108
 109
 110
 111
 112
 113
 114
 115
 116
 117
 118
 119
 120
 121
 122
 123
 124
 125
 126
 127
 128
 129
 130
 131
 132
 133
 134
 135
 136
 137
 138
 139
 140
 141
 142
 143
 144
 145
 146
 147
 148
 149
 150
 151
 152
 153
 154
 155
 156
 157
 158
 159
 160
 161
 162
 163
 164
 165
 166
 167
 168
 169
 170
 171
 172
 173
 174
 175
 176
 177
 178
 179
 180
 181
 182
 183
 184
 185
 186
 187
 188
 189
 190
 191
 192
 193
 194
 195
 196
 197
 198
 199
 200
 201
 202
 203
 204
 205
 206
 207
 208
 209
 210
 211
 212
 213
 214
 215
 216
 217
 218
 219
 220
 221
 222
 223
 224
 225
 226
 227
 228
 229
 230
 231
 232
 233
 234
 235
 236
 237
 238
 239
 240
 241
 242
 243
 244
 245
 246
 247
 248
 249
 250
 251
 252
 253
 254
 255
 256
 257
 258
 259
 260
 261
 262
 263
 264
 265
 266
 267
 268
 269
 270
 271
 272
 273
 274
 275
 276
 277
 278
 279
 280
 281
 282
 283
 284
 285
 286
 287
 288
 289
 290
 291
 292
 293
 294
 295
 296
 297
 298
 299
 300
 301
 302
 303
 304
 305
 306
 307
 308
 309
 310
 311
 312
 313
 314
 315
 316
 317
 318
 319
 320
 321
 322
 323
 324
 325
 326
 327
 328
 329
 330
 331
 332
 333
 334
 335
 336
 337
 338
 339
 340
 341
 342
 343
 344
 345
 346
 347
 348
 349
 350
 351
 352
 353
 354
 355
 356
 357
 358
 359
 360
 361
 362
 363
 364
 365
 366
 367
 368
 369
 370
 371
 372
 373
 374
 375
 376
 377
 378
 379
 380
 381
 382
 383
 384
 385
 386
 387
 388
 389
 390
 391
 392
 393
 394
 395
 396
 397
 398
 399
 400
 401
 402
 403
 404
 405
 406
 407
 408
 409
 410
 411
 412
 413
 414
 415
 416
 417
 418
 419
 420
 421
 422
 423
 424
 425
 426
 427
 428
 429
 430
 431
 432
 433
 434
 435
 436
 437
 438
 439
 440
 441
 442
 443
 444
 445
 446
 447
 448
 449
 450
 451
 452
 453
 454
 455
 456
 457
 458
 459
 460
 461
 462
 463
 464
 465
 466
 467
 468
 469
 470
 471
 472
 473
 474
 475
 476
 477
 478
 479
 480
 481
 482
 483
 484
 485
 486
 487
 488
 489
 490
 491
 492
 493
 494
 495
 496
 497
 498
 499
 500
 501
 502
 503
 504
 505
 506
 507
 508
 509
 510
 511
 512
 513
 514
 515
 516
 517
 518
 519
 520
 521
 522
 523
 524
 525

1 GAGCCGAGGTCGGGATGATCCTGAAGGAGGCCGTAAGGGAGTCCAGAGAGAACTT 60

61 CTGGAATTCGCCAAAAAGCTAAAAAAATGAGAAGAAGAAAGAAACCACCTGTCAG 120

b
61 CTGGAATGGCCAAAAAGTAAATAAAGAAGAAGACCAACCTCCTCGTTTC

121 CAGCTTTTGCATTCTTTCCTCCGTAATTCACCGGGG-
-----GCTCCTGTCCTCCACCGCACG

101
102
103
104
105
106
107
108
109
110
111
112
113
114
115
116
117
118
119
120
121
122
123
124
125
126
127
128
129
130
131
132
133
134
135
136
137
138
139
140
141
142
143
144
145
146
147
148
149
150
151
152
153
154
155
156
157
158
159
160
161
162
163
164
165
166
167
168
169
170
171
172
173
174
175
176
177
178
179
180
181
182
183
184
185
186
187
188
189
190
191
192
193
194
195
196
197
198
199
200
201
202
203
204
205
206
207
208
209
210
211
212
213
214
215
216
217
218
219
220
221
222
223
224
225
226
227
228
229
230
231
232
233
234
235
236
237
238
239
240
241
242
243
244
245
246
247
248
249
250
251
252
253
254
255
256
257
258
259
260
261
262
263
264
265
266
267
268
269
270
271
272
273
274
275
276
277
278
279
280
281
282
283
284
285
286
287
288
289
290
291
292
293
294
295
296
297
298
299
300
301
302
303
304
305
306
307
308
309
310
311
312
313
314
315
316
317
318
319
320
321
322
323
324
325
326
327
328
329
330
331
332
333
334
335
336
337
338
339
340
341
342
343
344
345
346
347
348
349
350
351
352
353
354
355
356
357
358
359
360
361
362
363
364
365
366
367
368
369
370
371
372
373
374
375
376
377
378
379
380
381
382
383
384
385
386
387
388
389
390
391
392
393
394
395
396
397
398
399
400
401
402
403
404
405
406
407
408
409
410
411
412
413
414
415
416
417
418
419
420
421
422
423
424
425
426
427
428
429
430
431
432
433
434
435
436
437
438
439
440
441
442
443
444
445
446
447
448
449
450
451
452
453
454
455
456
457
458
459
460
461
462
463
464
465
466
467
468
469
470
471
472
473
474
475
476
477
478
479
480
481
482
483
484
485
486
487
488
489
490
491
492
493
494
495
496
497
498
499
500
501
502
503
504
505
506
507
508
509
510
511
512
513
514
515
516
517
518
519
520
521
522
523
524
525
526
527
528
529
530
531
532
533
534
535
536
537
538
539
540
541
542
543
544
545
546
547
548
549
550
551
552
553
554
555
556
557
558
559
560
561
562
563
564
565
566
567
568
569
570
571
572
573
574
575
576
577
578
579
580
581
582
583
584
585
586
587
588
589
590
591
592
593
594
595
596
597
598
599
600
601
602
603
604
605
606
607
608
609
610
611
612
613
614
615
616
617
618
619
620
621
622
623
624
625
626
627
628
629
630
631
632
633
634
635
636
637
638
639
640
641
642
643
644
645
646
647
648
649
650
651
652
653
654
655
656
657
658
659
660
661
662
663
664
665
666
667
668
669
670
671
672
673
674
675
676
677
678
679
680
681
682
683
684
685
686
687
688
689
690
691
692
693
694
695
696
697
698
699
700
701
702
703
704
705
706
707
708
709
710
711
712
713
714
715
716
717
718
719
720
721
722
723
724
725
726
727
728
729
730
731
732
733
734
735
736
737
738
739
740
741
742
743
744
745
746
747
748
749
750
751
752
753
754
755
756
757
758
759
760
761
762
763
764
765
766
767
768
769
770
771
772
773
774
775
776
777
778
779
780
781
782
783
784
785
786
787
788
789
790
791
792
793
794
795
796
797
798
799
800
801
802
803
804
805
806
807
808
809
810
811
812
813
814
815
816
817
818
819
820
821
822
823
824
825
826
827
828
829
830
831
832
833
834
835
836
837
838
839
840
841
842
843
844
845
846
847
848
849
850
851
852
853
854
855
856
857
858
859
860
861
862
863
864
865
866
867
868
869
870
871
872
873
874
875
876
877
878
879
880
881
882
883
884
885
886
887
888
889
890
891
892
893
894
895
896
897
898
899
900
901
902
903
904
905
906
907
908
909
910
911
912
913
914
915
916
917
918
919

121 CACGTTTGCATGTTTCGCTATTCAAATGGCTTGATAGCTTGATATATGTTGGTGGGAC 180

181 AATGGCTGCATCATCCATGGAGCTGCACCTCCCTTCATGATGCTGCTTTTGGAAACAT 240

181 AATGGCTGCATCATCCATGGAGCTGCACCTCCCTCATGATGCTGTTTTCGAAACCG 340

241 GACAGATAGCTTTGCAAAATGCAGCATTTCACAATAAATCTTGGTGTG
-----CCCGCCTTAATGGGACCA

|||||GCGAGTGGCTTTCCAGTTATATTAA 300
|||||GCGAGTGGCTTTCCAGTTATATTAA 300

241 GCGAGTAGCTTTGCAANIGCAGGATTTTCAGAACAACAACTTTCCAGTTATAATTAA 300

301 TGAAGTATTACGAACAATACACACACATTTCATCAACCATCTGGAGGAGGAATGACCAC 360

301 TGAAGTATTACGAACATACACACATTTCATCACCATCTGGAGGAGGAATGACCAC 360

361 GATGCGCTATTATTACAGTGGGATCGGTACTGCCCTCCCTCCTCCCTCCCTTC

[illegible]

CGGCTGGTGGCTGCTTACATCCAGGT 420

421 TTTCA TTCTGTGCTGCGCAGCAGGAAGACAGATCTCAAATTAGAAAACAATTTTTTCA 480

421 TTCA TTC TGT GGT GCG CAG CAG A GAC AGA TACT CAA ATT GAG AAA CCA TTT TTT TCA 480

481 TGCTATCATCGGACGAGATTGGTGTGTTGACCTGCATTCACCTTCCCAAGCCTT

|||||
|||GCTATTCATTGGCGACAGCAGCAAAmccacgacggcgccttgaatgagcaaaatttttttt
|||
481 |||TGCCTATTCATTGGCGACAGCAGCAAAmccacgacggcgccttgaatgagcaaaatttttttt
|||

...:GACGTCATGACGTTGGGAGCTTAACAC 540

341 CCGGCTCACAGACGATGTCCTCCAAATCAATGAGCAATTGGCGACAAATTGGAATGTT 600

541 CCGGCTCACAGACGATGTCCTCCAAATCATGAGCAATTGGCGACAAAATTGGCAATGTT 600

601 CTCTTCACTCAATAGCAACATTTTTCACCGGTTTATATAGTGGGGTTTACACCCCTCCCTCCG

601 CTTTCATGACATAGCAACATTCTTCTTGCCCGGGGGTGTGAGGGA 860

.....GCGGGCTTAAAGTCGGTTACACGTGTTGGA 660

501 GCACCCCTTGATTTTGGCCATCAGCCCTGTTCTTGGACTTTCAGCCGCCATCTGGCC 720

661 GCTAACCCCTGTGATTTTGGCCATCAGCCCTGTTCTTGGACTTTCAGCCGCCATCTGGGC 720

721 AAGATACATCTTCATTACTGATAAGAACTCTTGCCCTATGCAAAAGCTGAACT 780

[illegible][illegible]

101 ACCGAGGAGAGCTCTACGACGCAATCAGAACTGTGATTGCCCTTGGAGGACAAAAGAAGA 840

OY	841	ACTTGAAGGTAGACACAAAATTTAGAGAGCTAAAGAAATTTGGATATTAAGAAAGCTAT	900
Db	841	ACTTGAAGGTAGACACAAAATTTAGAGAGCTAAAGAAATTTGGATATTAAGAAAGCTAT	900
OY	901	CACGGCCAAACATTTCTATGTGGCGGCTTCTTATGTATCTATGATCATATATCTCTTGCC	960
Db	901	CACGGCCAAACATTTCTATGTGGCGGCTTCTTATGTATCTATGATCATATATCTCTTGCC	960
OY	961	TTTTGTGATGGGACCTCTTGCTCTCTCCAGTGAATATCTAATTTGGCAACTACTCTAC	1020
Db	961	TTTTGTGATGGGACCTCTTGCTCTCTCCAGTGAATATCTAATTTGGCAACTACTCTAC	1020
OY	1021	TGCTCTCTTTCTGATTAATTTGGCGGCTTTAGTATTGGACAGGCAATCCCAAGCATTTGA	1080
Db	1021	TGCTCTCTTTCTGATTAATTTGGCGGCTTTAGTATTGGACAGGCAATCCCAAGCATTTGA	1080
OY	1081	AGCATTTTGCACAAAGCAGAGAGACAGCTTATGAATATCTCAAGTATTTGACATATAAC	1140
Db	1081	AGCATTTTGCACAAAGCAGAGAGACAGCTTATGAATATCTCAAGTATTTGACATATAAC	1140
OY	1141	AGCATTTGACAGCTATTGGAAGAGTGGACATTAACACAGATATATTTAAGGAATTTTGA	1200
Db	1141	AGCATTTGACAGCTATTGGAAGAGTGGACATTAACACAGATATATTTAAGGAATTTTGA	1200
OY	1201	ATTCAAAAATTTTCACTTCAATTAACCTTTCGAAAAGAAATTAAGATCTTAAAGGGTCT	1260
Db	1201	ATTCAAAAATTTTCACTTCAATTAACCTTTCGAAAAGAAATTAAGATCTTAAAGGGTCT	1260
OY	1261	CAACTGGAAGTTTCAGAGTGGGCAACAGCTGGCGCTGTTGGGAAACATGGGCTCGGGAA	1320
Db	1261	CAACTGGAAGTTTCAGAGTGGGCAACAGCTGGCGCTGTTGGGAAACATGGGCTCGGGAA	1320
OY	1321	GACGACAGCCGTGCAGTGTGAGAGGCGCTATATGACCCACAGATGGCATGTCTGAT	1380
Db	1321	GACGACAGCCGTGCAGTGTGAGAGGCGCTATATGACCCACAGATGGCATGTCTGAT	1380
OY	1381	TGATGACAGGACATTTAGACCACTAAATTTAAGGCATCTTGGGAAATTAATCTGCTGGT	1440
Db	1381	TGATGACAGGACATTTAGACCACTAAATTTAAGGCATCTTGGGAAATTAATCTGCTGGT	1440
OY	1441	GAGTCAGAGCGCTGTGTTTCCACACAGATAGCTGAAAACATTCGCTATGGCCGCGA	1500
Db	1441	GAGTCAGAGCGCTGTGTTTCCACACAGATAGCTGAAAACATTCGCTATGGCCGCGA	1500
OY	1501	AAATGTCCACATGATGAGATTGAGAAAGCGTTAAGGAAGCCAAATGCGTATGATTTAT	1560
Db	1501	AAATGTCCACATGATGAGATTGAGAAAGCGTTAAGGAAGCCAAATGCGTATGATTTAT	1560
OY	1561	CATGAATCTACCTATATAATTTGCACTCTGTTTGAGAGAGAGGGGCCAGCTGAGTGG	1620
Db	1561	CATGAATCTACCTATATAATTTGCACTCTGTTTGAGAGAGAGGGGCCAGCTGAGTGG	1620
OY	1621	TGGACAGAAACAGAGAAATCGCCATTTGCTGGGGCCCTGTTGCGAACCCCAAGATTTCTT	1680
Db	1621	TGGACAGAAACAGAGAAATCGCCATTTGCTGGGGCCCTGTTGCGAACCCCAAGATTTCTT	1680
OY	1681	GCTGATGAGGCAAGCTCCTGAGACTTGGAACCTGAAATGTAAGCTGTTCAAGTGGCCCT	1740
Db	1681	GCTGATGAGGCAAGCTCCTGAGACTTGGAACCTGAAATGTAAGCTGTTCAAGTGGCCCT	1740
OY	1741	GGATTAAGGCCAAGAAAGCCCGGACATCTGTATGATCTCATGCTTGTCTTACAGTTTG	1800
Db	1741	GGATTAAGGCCAAGAAAGCCCGGACATCTGTATGATCTCATGCTTGTCTTACAGTTTG	1800
OY	1801	TAAATGCGATGTCATTTGCTGTTTGTATGATGAGTCAATTTGTGGAGAAAGGAATTCATGA	1860
Db	1801	TAAATGCGATGTCATTTGCTGTTTGTATGATGAGTCAATTTGTGGAGAAAGGAATTCATGA	1860
OY	1861	TGAACATCATGAAGAGAGGCAATTTCAACTGTGTCAAAATGCAGACAAGAGGAAA	1920
Db	1861	TGAACATCATGAAGAGAGGCAATTTCAACTGTGTCAAAATGTGTCAAAATGCAGACAAGAGGAAA	1920

QY	1921	TGAATGAGTTAGAAAATGCAATGCTGTGATCCAAAAGTGAAGTGATGCCCTTGGAAAT	1980
Db	1921	TGAAATTCAGTTAAGAAAATGCATGCTGTAATCCAAAAGAAAGTGAAGCTCTGGAAAT	1980
QY	1981	GTCTCCAAAGATTCAGGTCCTCATTTAATTAAGAAAGATTCAACTCCGAGAGATATACA	2040
Db	1981	GTCTCCAAAAGATTCAGGTCCTCATTTAATTAAGAAAGATTCAACTCCGAGAGATATACA	2040
QY	2041	TGCACCAACAAGCCACAGACAGAAAGCTTGGTACAAAAGAGACTTGAATGAGAAATGACC	2100
Db	2041	TGCACCAACAAGCCACAGACAGAAAGCTTGGTACAAAAGAGACTTGAATGAGAAATGACC	2100
QY	2101	TCCAGTTCTCTCTGAGAGATTTTGAAGCTGTAACCTCAACTGAAATGGCCTATTTTGTGT	2160
Db	2101	TCCAGTTCTCTCTGAGAGATTTTGAAGCTTGAAGCTGTAACCTCAACTGAAATGGCCTATTTTGTGT	2160
QY	2161	TGGTATATTTTGTGTAATTAATAACGAGGCTGCACACAGCATTTTTCATTAATATTTTC	2220
Db	2161	TGGTATATTTTGTGTAATTAATAACGAGGCTGCACACAGCATTTTTCATTAATATTTTC	2220
QY	2221	AAGGATTAAGGATCTTTACCCGAGATGAGGATCTCAACAAACAAAGACAGAAATAGTAA	2280
Db	2221	AAGGATTAAGGATCTTTACCCGAGATGAGGATCTCAACAAACAAAGACAGAAATAGTAA	2280
QY	2281	CAGTTTTCGTATGTGTTCTAGTCCCTTGGAAATATTTCTTTTATTAATTTTCCCTCCA	2340
Db	2281	CAGTTTTCGTATGTGTTCTAGTCCCTTGGAAATATTTCTTTTATTAATATTTTCCCTCCA	2340
QY	2341	GGGCTTCACATTTGGCAAAAGCTGGGGAGATGCTCACTAAGCGGCTTCATACATGTTTT	2400
Db	2341	GGGCTTCACATTTGGCAAAAGCTGGGGAGATGCTCACTAAGCGGCTTCATACATGTTTT	2400
QY	2401	CAGATTCATCTAGACAGAGATGTACGTGTGTTTATATGACCTTAATAAACACACCTGGAGC	2460
Db	2401	CAGATTCATCTAGACAGAGATGTACGTGTGTTTATATGACCTTAATAAACACACCTGGAGC	2460
QY	2461	ATTGACAACCAAGCTTGGCAATGATGGGGCTCAAGTTAAAGGGGCTATATAGTTCCAGGCT	2520
Db	2461	ATTGACAACCAAGCTTGGCAATGATGGGGCTCAAGTTAAAGGGGCTATATAGTTCCAGGCT	2520
QY	2521	TGCTGTCTATACCCAGAAATATAGCAAAATCTTGGACAGCATTAATATATNCTTATATCA	2580
Db	2521	TGCTGTCTATACCCAGAAATATAGCAAAATCTTGGACAGCATTAATATATNCTTATATCA	2580
QY	2581	TGGTTGGCAATTAACACTTTTACCTTATAGCAATGTACCCATATTTGCATATAGCAGAGT	2640
Db	2581	TGGTTGGCAATTAACACTTTTACCTTATAGCAATGTACCCATATTTGCATATAGCAGAGT	2640
QY	2641	TGTTAAATGAATAATGTGTCTGTGACAAAGCAGTGAAGATTAAGAAAGAGTATAAGAGAGC	2700
Db	2641	TGTTAAATGAATAATGTGTCTGTGACAAAGCAGTGAAGATTAAGAAAGAGTATAAGAGAGC	2700
QY	2701	TGGGAAAGTTGCTACAGAAAGCATGGAAGCTTCCGAACCTGTCTTGTGACTCGGGA	2760
Db	2701	TGGGAAAGTTGCTACAGAAAGCATGGAAGCTTCCGAACCTGTCTTGTGACTCGGGA	2760
QY	2761	GCAGAGTTTGAATACATGATATGACAGAGTTTGCAGATACCATTAACGAAACCTTTGAG	2820
Db	2761	GCAGAGTTTGAATACATGATATGACAGAGTTTGCAGATACCATTAACGAAACCTTTGAG	2820
QY	2821	GAAAGCACATCTTGGGGCTGCAATTTCTATCCACAGGCAGATGATATTTTTCCTA	2880
Db	2821	GAAAGCACATCTTGGGGCTGCAATTTCTATCCACAGGCAGATGATATTTTTCCTA	2880
QY	2881	TGCTGGCTGTTCCGGTTTGGTTCCTACTTGTGTGCAATAGCTTCAATGAACCTTTACAGA	2940
Db	2881	TGCTGGCTGTTCCGGTTTGGTTCCTACTTGTGTGCAATAGCTTCAATGAACCTTTACAGA	2940
QY	2941	TGTTCTTTTGGTATCTCAGCTATTTCTTGTGTGCATGAGGAGGGGCAAGTCACTTC	3000
Db	2941	TGTTCTTTTGGTATCTCAGCTATTTCTTGTGTGTGCATGAGGAGGGGCAAGTCACTTC	3000
QY	3001	ATTGTGCTCGACATATCCAAAAGCAAAAGTATACGACGCCACGTCATCATCATTTGA	3060

XX Sequence 4317 BP; 1293 A; 844 C; 1019 G; 1161 T; 0 other;
SQ

Score 4221.8; DB 22; Length 4317;

Query Match	50.7%	Score	100
Best Local Similarity	99.5%	Pred. No. 0;	
Matches 4245; Conservative		0; Mismatches	17; Indels 3; Gaps 1.

OY	3	AGGCCAAGGTCGGATGGATGATCTTGAAAGAGCCCGTAAAGGGAGTGCAGAGAAACAATCTTC	62
Db	56	AGCCCGAGGTGCAGATGGATCTCTGAAGAGGCGGTAAAGGGAGTGCACAGAAAGAACTTCT	115
OY	63	GGAANAATGGGCAAAAAAAGTATAAAAAATAGAGAAAGAAAGAAACAACCTGTACAGA	122
Db	116	GGAAAAATGGGCAAAAAAAGTAAAAAA--GAGAGAAAGAAAAAGAAACAACCTGTACAGA	172
OY	123	CGTTTGCATAGTTTGCCTATTCCAAATTGGCTGTATAGGTTGTATATGTTTGGTGGGACAA	182
Db	173	CGTTTGCATAGTTTGCCTATTCCAAATTGGCTGTATAGGTTGTATATGTTTGGTGGGACAA	232
OY	183	TGGCTGCCATCATCATGAGAGCTGACCTCCCTCTCAAGATGCGTGGTTTTGGAAAAATGA	242
Db	233	TGGCGGCATCATCATGAGAGCTGACCCCTCCATCAAGATGCGTTTTTGGAAAAACATGA	292
OY	243	CAGATAGCTTTGCAAATGACAGGAATTTCAAGAAACAACAATTTTCCAGTATATATTAATG	302
Db	293	CAGATAGCTTTGCAAATGACAGGAATTTCAAGAAACAACAATTTTCCAGTATATATTAATG	352
OY	303	AAAGATTACGAAACATACATCAACAATTTTCATCACCATCTGGAGAGAGAAATGACACAGT	362
Db	353	AAAGATTACGAAACATACATCAACAATTTTCATCACCATCTGGAGAGAGAAATGACACAGT	412
OY	363	ATGCCATTATTTACAGTGGGATGGGTGCGGGGTGCTGGTGGCTTATCATCCAGGTTT	422
Db	413	ATGCCATTATTTACAGTGGGATGGGTGCGGGGTGCTGGTGGCTTATCATCCAGGTTT	472
OY	423	CATTCGGGAGCCCTGGCAGCAGGAAAGACAGATACCTCAAAATTAAGAAACAATTTTTCATG	482
Db	473	CATTCGGGAGCCCTGGCAGCAGGAAAGACAGTACCTCAAAATTAAGAAACAATTTTTCATG	532
OY	483	CTATCATGCGACAGAGATTGGCTGGTTGAGTGCATGAGAGCTTGGGAGCTTTACACCC	542
Db	533	CTATCATGCGACAGAGATTGGCTGGTTGAGTGCATGAGAGCTTGGGAGCTTTACACCC	592
OY	543	GGCTCACAGACAGTGTCTCCAAATCAATGAAGAAATGGCGACAAAATTTGGAATGTCT	602
Db	593	GGCTCACAGACAGTGTCTCCAAATCAATGAAGAAATTTGGCACAAATTTGGAATGTCT	652
OY	603	TTCACTCAATAGCAACATTTTTCACCGGTTTATAGTGGGGTTTACAGTGGTTGGAAAG	662
Db	653	TTCACTCAATAGCAACATTTTTCACCGGTTTATAGTGGGGTTTACACTGTGTTGGAAGC	712
OY	663	TAACCTTTTGATTTTGGCCATCAGCCCTGTCTTGAGACTTTCAGCCGCACTTGGGCAA	722
Db	713	TAACCTTTTGATTTTGGCCATCAGCCCTGTCTTGAGACTTTCAGCCGCACTTGGGCAA	772
OY	723	AGATACATCTTCATTTACTGTATTAAGAATCTTGGCTATGTGCAAAAGGTGGAGCAGAG	782
Db	773	AGATACATCTTCATTTACTGTATTAAGAATCTTGGCTATGTGCAAAAGGTGGAGCAGAG	832
OY	783	CTGAAGAAGCTTTCAGCAGCAATCAGACGTGATGGCTTTGGAGACAAAAGAAAGAC	842
Db	833	CTGAAGAAGCTTTCAGCAGCAATCAGACTGTGATGGCTTTGGAGACAAAAGAAAGAC	892
OY	843	TTGAAGAAGTACAAACAAAATTTTGAAGAAGCTTAAAGAAATTTGGATTAAGAAAGCTATCA	902
Db	893	TTGAAGAAGTACAAACAAAATTTTGAAGAAGCTTAAAGAAATTTGGATTAAGAAAGCTATCA	952
OY	903	CGGCCAACATTTCTATTGGTGGCGCTTCTTATTTGATCTATGACATATATGCTTGCGCTT	962
Db	953	CGGCCAACATTTCTATTGGTGGCGCTTCTTATTTATCTATGACATATATGCTTGCGCTT	1012
OY	963	TCGCGATAGGGGACCCCTGGTGGCTCTCCAGTGAATATTTCTATGGACAATCTCACTG	1022

Db	1013	TCTGGTAATGGGACCTCCTTGCTGCTCCAGTAGTAATATACATTTGGACAGATACACTG	1072
Oy	1023	TCTTCTTTTCGTATTAATATGGGGCTTTAGATTGGACAGGCATCCCAACATTTGAAG	1082
Db	1073	TCTTCTTTTCGTATTAATATGGGCTTTTAGTTGGACAGGCATCCCAACATTTGAAG	1132
Oy	1083	CATTTGCAAAAGCAAGAGAGAGACCTTATGAAATCTTCAGATTAATTTGCAATTAACCA	1142
Db	1133	CATTTGCAAAAGCAAGAGAGACACTTATGAAATCTTCAGATTAATTTGCAATTAACCA	1192
Oy	1143	GCATTTGACAGCATTTGCAAGAGTGGACATTAACAGATTAATTTAAGGAAATTTGAAAT	1202
Db	1193	GCATTTGACAGCATTTGCAAGAGTGGACATTAACAGATTAATTTAAGGAAATTTGAAAT	1252
Oy	1203	TCAAAATGTTCATTCAGTTACCTTCTCGAAAAAGAAAGTTAAGTCTTAAAGGGCTCA	1262
Db	1253	TCAAAATGTTCATTCAGTTACCTTCTCGAAAAAGAAAGTTAAGTCTTAAAGGGCTCA	1312
Oy	1263	ACCTGAAGTTCCAGATGGGGGACAGAGTGGGCTGTGTGGACAGTGGCTCGGGAGAG	1322
Db	1313	ACCTGAAGTTCCAGATGGGGGACAGAGTGGGCTGTGTGGACAGTGGCTCGGGAGAG	1372
Oy	1323	GCACAGCCGTGAGCTGATGCAAGAGCTTATGACCCCAAGATGGCATGGTCTGTATTG	1382
Db	1373	GCACAGCCGTGAGCTGATGCAAGAGCTTATGACCCCAAGATGGCATGGTCTGTATTG	1432
Oy	1383	ATGACAGGACATTTAGGACATTAATATGTAAGGCATCTTCGGGAAATTACTGGTGTGTA	1442
Db	1433	ATGACAGGACATTTAGGACATTAATATGTAAGGCATCTTCGGGAAATTACTGGTGTGTA	1492
Oy	1443	GTCAGGAGCCGTGTGTGTGTGCCCACAGAGTATAGTGAATTTGGCTATGGCCGGGAAA	1502
Db	1493	GTCAGGAGCCGTGTGTGTGTGCCCACAGAGTATAGTGAATTTGGCTATGGCCGGGAAA	1552
Oy	1503	ATGTCACCATGATGAGATTGCAAAAGCTGTATAGGAAGCCCATCCATGATTTTATCA	1562
Db	1553	ATGTCACCATGATGAGATTGCAAAAGCTGTATAGGAAGCCCATCCATGATTTTATCA	1612
Oy	1563	TGAAACTCTCAATTAATTTGACACTGTGGTTGGAAGAGAGGGGCCACGCTGAGATGGTG	1622
Db	1613	TGAAACTCTCAATTAATTTGACACTGTGGTTGGAAGAGAGGGGCCACGCTGAGATGGTG	1672
Oy	1623	GACAGAAACAGAGATCGCCCTTCTGCTGGGGCCCTGGTGCACACCCCAAGATTTCTTGC	1682
Db	1673	GACAGAAACAGAGATCGCCCTTCTGCTGGGGCCCTGGTGCACACCCCAAGATTTCTTGC	1732
Oy	1683	TGGATGAGGACACGTCAGCTCTGACACTGCACTGAAGTGAAGCAGTGGTCCAGTGGCCCTGG	1742
Db	1733	TGGATGAGGACACGTCAGCTCTGACACTGCACTGAAGTGAAGCAGTGGTCCAGTGGCCCTGG	1792
Oy	1743	ATAAGCCAGAAAGGCCGAGCTACCATTTGATAGCTCATCTGTTGCTTCAAGTTGTA	1802
Db	1793	ATAAGCCAGAAAGGCCGAGCTACCATTTGATAGCTCATCTGTTGCTTCAAGTTGTA	1852
Oy	1803	ATGCGCATGTCATGTGCTGTTTGTATGATGAGATCATTTGGAGAAAGGAAATCATGATG	1862
Db	1853	ATGCGCATGTCATGTGCTGTTTGTATGATGAGATCATTTGGAGAAAGGAAATCATGATG	1912
Oy	1863	AACTCATGAAGAAGAGGSCATTTTACTTCAAACTGTACATCCAGACAGAGGAATG	1922
Db	1913	AACTCATGAAGAAGAGGSCATTTTACTTCAAACTGTACATCCAGACAGAGGAATG	1972
Oy	1923	AAATGTGTTAGAAAAATGCGCACTGGTGAATCCAAAAGTGAAGTATGATCCCTTGGAAATGT	1982
Db	1973	AAATGTGTTAGAAAAATGCGCACTGGTGAATCCAAAAGTGAAGTATGATCCCTTGGAAATGT	2032
Oy	1983	CTCCAAAAGATTCAGGGTCCAGTTTAATAAAAAGAAAGATCACTGCGAGGAGTATACATG	2042
Db	2033	CTCCAAAAGATTCAGGGTCCAGTTTAATAAAAAGAAAGATCACTGCGAGGAGTATACATG	2092
Oy	2043	CACCAACAAGGCAAGACAGAAAGCTTGTATCAAAAGAGAGCTTGAATGAGATGTACTCT	2102
Db	2093	CACCAACAAGGCAAGACAGAAAGCTTGTGTCAAAAAGAGAGCTTGAATGAGATGTACTCT	2152

Db	1013	TCTGGTAATGGGACCTCCTTGCTGCTCCAGTAGTAATATACATTTGGACAGATACACTG	1072
Oy	1023	TCTTCTTTTCGTATTAATATGGGGCTTTAGATTGGACAGGCATCCCAACATTTGAAG	1082
Db	1073	TCTTCTTTTCGTATTAATATGGGGCTTTAGATTGGACAGGCATCCCAACATTTGAAG	1132
Oy	1083	CATTTGCAAAAGCAAGAGAGAGACCTTATGAAATCTTCAGATTAATTTGCAATTAACCA	1142
Db	1133	CATTTGCAAAAGCAAGAGAGAGACCTTATGAAATCTTCAGATTAATTTGCAATTAACCA	1192
Oy	1143	GCATTTGACAGCATTTGCAAGAGTGGACATTAACAGATTAATTTAAGGAAATTTGAAAT	1202
Db	1193	GCATTTGACAGCATTTGCAAGAGTGGACATTAACAGATTAATTTAAGGAAATTTGAAAT	1252
Oy	1203	TCAAAATGTTCATTCAGTTACCTTCTCGAAAAAGAAAGTTAAGTCTTAAAGGGCTCA	1262
Db	1253	TCAAAATGTTCATTCAGTTACCTTCTCGAAAAAGAAAGTTAAGTCTTAAAGGGCTCA	1312
Oy	1263	ACCTGAAGGTTCCAGATGGGGGACAGAGTGGGCTGTGTGGACAGTGGCTCGGGAGA	1322
Db	1313	ACCTGAAGGTTCCAGATGGGGGACAGAGTGGGCTGTGTGGACAGTGGCTCGGGAGA	1372
Oy	1323	GCACGACCGTGCAGCTGATGCGAGAGGCTTATGACCCCAAGATGGCATGGTCTGATTTG	1382
Db	1373	GCACGACCGTGCAGCTGATGCGAGAGGCTTATGACCCCAAGATGGCATGGTCTGATTTG	1432
Oy	1383	ATGACAGAGCATTTAGGACATTAATATGAAGCATCTTTCGGAAATTAATCTGGTGTGTA	1442
Db	1433	ATGACAGAGCATTTAGGACATTAATATGAAGCATCTTTCGGAAATTAATCTGGTGTGTA	1492
Oy	1443	GTCAGGAGCCGTGTGTGTGTGTCACCACAGANTAGCTGAAAACTTGTGCTATGCGCCGAAA	1502
Db	1493	GTCAGGAGCCGTGTGTGTGTGTCACCACAGANTAGCTGAAAACTTGTGCTATGCGCCGAAA	1552
Oy	1503	ATGTCCACCATGATGAGATTGCAAAAGCTGTATGAAGAACCAATCCATGATTTTATCA	1562
Db	1553	ATGTCCACCATGATGAGATTGCAAAAGCTGTATGAAGAACCAATCCATGATTTTATCA	1612
Oy	1563	TGAAACTCTCAATTAATTTGACACTGTGGTTGGAAGAGAGGGGCCACGCTGAGATGGTG	1622
Db	1613	TGAAACTCTCAATTAATTTGACACTGTGGTTGGAAGAGAGGGGCCACGCTGAGATGGTG	1672
Oy	1623	GACAGAAAACAGAAATCGCCCTTCTCTGGGGCCCTGGTGCACACCCCAAGATTTCTTGC	1682
Db	1673	GACAGAAAACAGAAATCGCCCTTCTCTGGGGCCCTGGTGCACACCCCAAGATTTCTTGC	1732
Oy	1683	TGGATGAGGCAACGTCAGCTCTGACACTGCACTGAAGTGAAGCAGTGGTCCAGTGGCCCTGG	1742
Db	1733	TGGATGAGGCAACGTCAGCTCTGACACTGCACTGAAGTGAAGCAGTGGTCCAGTGGCCCTGG	1792
Oy	1743	ATAAGCCAGAAAAGCCCGAGCTACCATTTGATAGCTCATCTGTTGCTCTACAGTTGTA	1802
Db	1793	ATAAGCCAGAAAAGCCCGAGCTACCATTTGATAGCTCATCTGTTGCTCTACAGTTGTA	1852
Oy	1803	ATGCCAGTGCATGTGCTGTTTGTATGATGAGATCATTTGGAGAAAGGAAATCATGATG	1862
Db	1853	ATGCCAGTGCATGTGCTGTTTGTATGATGAGATCATTTGGAGAAAGGAAATCATGATG	1912
Oy	1863	AACTCATGAAGAAGAGGCGATTTACTTCAAACTGTACATCCAGACAGAGGAATG	1922
Db	1913	AACTCATGAAGAAGAGGCGATTTACTTCAAACTGTACATCCAGACAGAGGAATG	1972
Oy	1923	AAATGAGTTAGAAAATGCGCACTGGTGAATCCAAAAGTGAAGTCAATCCCTTGGAAATGT	1982
Db	1973	AAATGAGTTAGAAAATGCGCACTGGTGAATCCAAAAGTGAAGTCAATCCCTTGGAAATGT	2032
Oy	1983	CTCCAAAAGATTCAGGGTCCAGTTTAATTAAGAAAGATCACTGCGAGGAGTATACATG	2042
Db	2033	CTCCAAAAGATTCAGGGTCCAGTTTAATTAAGAAAGATCACTGCGAGGAGTATACATG	2092
Oy	2043	CACCAACAAGGCAAGACAGAAAGCTTGTATCAAAAGAGAGCTTGAATGAGATGTACCTC	2102
Db	2093	CACCAACAAGGCAAGACAGAAAGCTTGTGTCAAAAAGAGAGCTTGAATGAGATGTCTCT	2152

```
OY 2103 CAGTTTCCTCTGAGAGATTCGAAAGCTGAACTCAAGTGAAGGCTTATTTTGGCTG 2162
|||||
Db 2153 CAGTTTCCTCTGAGAGATTCGAAAGCTGAACTCAAGTGAAGGCTTATTTTGGCTG 2212
|||||
OY 2163 GTATATTTTGTGCTATTTATTAAGAGGCGCGACACAGACATTTTCATTAATTTTCAA 2222
|||||
Db 2213 GTATATTTTGTGCTATTTATTAAGAGGCGCGACACAGACATTTTCATTAATTTTCAA 2272
|||||
OY 2223 GGATTTATAGAGATCTTTTACCAGATGAGATTCCTGAAACAAACAGACAGAAATAGTAA 2282
|||||
Db 2273 GGATTTATAGAGATCTTTTACCAGATGAGATTCCTGAAACAAACAGACAGAAATAGTAA 2332
|||||
OY 2283 TGTTTCTGTATTTGTTTCTAGTCTTGGAATTTATTTCTTTTATTTACATTTTCCGAGG 2342
|||||
Db 2333 TGTTTCTGTATTTGTTTCTAGTCTTGGAATTTATTTCTTTTATTTACATTTTCCGAGG 2392
|||||
OY 2343 GCTTCACATTTTGGCAAGCTGGGGAGATCCCTCACTAAAGCGGCTTGATATCATGTTTCA 2402
|||||
Db 2393 GCTTCACATTTTGGCAAGCTGGGGAGATCCCTCACTAAAGCGGCTTGATATCATGTTTCA 2452
|||||
OY 2403 GATCCATGCTGAGACAGATGTCTAGCTGCTTGTATGACCTTAAACACACACTGGAGCAT 2462
|||||
Db 2453 GATCCATGCTGAGACAGATGTCTAGCTGCTTGTATGACCTTAAACACACACTGGAGCAT 2512
|||||
OY 2463 TGACAAACAGGCTTGGCAATGATGGGCTCAAGTTAAAGGGCTATAGGTTCCAGGCTTG 2522
|||||
Db 2513 TGACAAACAGGCTTGGCAATGATGGGCTCAAGTTAAAGGGCTATAGGTTCCAGGCTTG 2572
|||||
OY 2523 CTGTCAATTTCCAGAAATATAGCAATCTTGGGACAGGCAATTTATATCCTTAATCATG 2582
|||||
Db 2573 CTGTCAATTTCCAGAAATATAGCAATCTTGGGACAGGCAATTTATATCCTTAATCATG 2632
|||||
OY 2583 GTTGGCAATTTACACTTTTACTTCTTACATTTGTACCATCATTTGATAGAGAGTTG 2642
|||||
Db 2633 GTTGGCAATTTACACTTTTACTTCTTACATTTGTACCATCATTTGATAGAGAGTTG 2692
|||||
OY 2643 TTGAATGAAATGTTGTCTGACAGACAGCTGAAAGATTAAGAAAGCTAGAAAGAGCTG 2702
|||||
Db 2693 TTGAATGAAATGTTGTCTGACAGACAGCTGAAAGATTAAGAAAGCTAGAAAGAGCTG 2752
|||||
OY 2703 GGAAGATTTGCTACAGAAAGCATTCGAAACTCGAACTGTCTTTCTTTGACTGGGAGC 2762
|||||
Db 2753 GGAAGATTTGCTACAGAAAGCATTCGAAACTCGAACTGTCTTTCTTTGACTGGGAGC 2812
|||||
OY 2763 AGAAGTTGAATACATGTATGACAGAGTTGCAAGTACCAACAGAACTCTTTGAGGA 2822
|||||
Db 2813 AGAAGTTGAATACATGTATGACAGAGTTGCAAGTACCAACAGAACTCTTTGAGGA 2872
|||||
OY 2823 AAGCACACATCTCGGGGCTCATTTTCTATCAACCAGGCAATGATATTTTCTATG 2882
|||||
Db 2873 AAGCACACATCTCGGGGCTCATTTTCTATCAACCAGGCAATGATATTTTCTATG 2932
|||||
OY 2883 CTGGCTGTTTTCGGGTTTGGTGCTTACCTTGTGGTCAAAATGATGTAAGAACTTACAGATG 2942
|||||
Db 2933 CTGGCTGTTTTCGGGTTTGGTGCTTACCTTGTGGTCAAAATGATGTAAGAACTTACAGATG 2992
|||||
OY 2943 TTTCTTTTGGTATTCTCAGCTATTTGCTTGGTGCTATGGCAGTGGGCGAGTCACTCAT 3002
|||||
Db 2993 TTTCTTTTGGTATTCTCAGCTATTTGCTTGGTGCTATGGCAGTGGGCGAGTCACTCAT 3052
|||||
OY 3003 TTGCTCTGACTATGCGCAAGCGCAAGATATCAGAGCGCCACGTCAATCATATGATGAA 3062
|||||
Db 3053 TTGCTCTGACTATGCGCAAGCGCGCAAGATATCAGAGCGCCACGTCAATCATATGATGAA 3112
|||||
OY 3063 AAAACCCCTGATTTGACAGCTACAGCCCTCAGCGGCGCAAGCAAAATGCTTGAAGAA 3122
|||||
Db 3113 AAAACCCCTGATTTGACAGCTACAGCCCTCAGCGGCGCAAGCAAAATGCTTGAAGAA 3172
|||||
OY 3123 ATGTGACATTTATAGAGCTGCTGTTCAACTATCCCACTGACCAAGACATCCCGCTGCC 3182
|||||
Db 3173 ATGTGACATTTATAGAGCTGCTGTTCAACTATCCCACTGACCAAGACATCCCGCTGCC 3232
|||||

OY 3183 AAGGCTGAGCCTCGAGGTGAAGAGGCGCCAGACGCTGGCCCTGTAAGTAGAGTGGCT 3242
|||||
Db 3233 AAGGCTGAGCCTCGAGGTGAAGAGGCGCCAGACGCTGGCCCTGTAAGTAGAGTGGCT 3292
|||||
OY 3243 GTGGGAAGGACACATTTGCTAGCTCTAGAGGCGCTTATATACCCCTTGGCTGGTTCA 3302
|||||
Db 3293 GTGGGAAGGACACATTTGCTAGCTCTAGAGGCGCTTATATACCCCTTGGCTGGTTCA 3352
|||||
OY 3303 TGCTAATTTATGCAAGAGATTAAGACCTGATGTCCAGTGGCTCCGACACACTGG 3362
|||||
Db 3353 TGCTAATTTATGCAAGAGATTAAGACCTGATGTCCAGTGGCTCCGACACACTGG 3412
|||||
OY 3363 GCATCTGTCTCAGAGGCCCATTCCTGTTTACTGACAGCATTTGCCAGAAATCTCATG 3422
|||||
Db 3413 GCATCTGTCTCAGAGGCCCATTCCTGTTTACTGACAGCATTTGCCAGAAATCTCATG 3472
|||||
OY 3423 GAGACAAAGCGGCGGTGATACATGAAGATTTATGAGGACAGCCAAAGAGGCGCAAGA 3482
|||||
Db 3473 GAGACAAAGCGGCGGTGATACATGAAGATTTATGAGGACAGCCAAAGAGGCGCAAGA 3532
|||||
OY 3483 TACACACTTCAATGAGACACTCCCTGAGAAATACAAACCGAGTTAGAGACAAAGAA 3542
|||||
Db 3533 TACACACTTCAATGAGACACTCCCTGAGAAATACAAACCGAGTTAGAGACAAAGAA 3592
|||||
OY 3543 CCCAGCTCTCTGTGGCCAGAAACAGCCCATTTGCCATTAAGCTGCGCTTTTGTAGACGC 3602
|||||
Db 3593 CCCAGCTCTCTGTGGCCAGAAACAGCCCATTTGCCATTAAGCTGCGCTTTTGTAGACGC 3652
|||||
OY 3603 CTGATATTTTGTGTTTGGATGAAGCTACATCAGCTGTGATGATGACAAAGTAAAGTTG 3662
|||||
Db 3653 CTGATATTTTGTGTTTGGATGAAGCTACATCAGCTGTGATGATGACAAAGTAAAGTTG 3712
|||||
OY 3663 TCCAAAGCCCTGGACAAAGCCAGAGAGGCGGACCTGCACTTGTATGCTGCCACGCT 3722
|||||
Db 3713 TCCAAAGCCCTGGACAAAGCCAGAGAGGCGGACCTGCACTTGTATGCTGCCACGCT 3772
|||||
OY 3723 TGTCCACATCCAGAAATCAGATTTAATAGTGTGTTTCAGATGGAAGGCAAGGAGC 3782
|||||
Db 3773 TGTCCACATCCAGAAATCAGATTTAATAGTGTGTTTCAGATGGAAGGCAAGGAGC 3832
|||||
OY 3783 ATGCGACATCAATCAAGCTGTGCGCCAGAAAGGCACTATTTTCCATGTGCTGCTCC 3842
|||||
Db 3833 ATGCGACATCAATCAAGCTGTGCGCTCAGAAAGCATATTTTCCATGATCAGTGTCC 3892
|||||
OY 3843 AGGCTGGACAAAGGCGCTAGTGAACCTGGCCATATGAGCTGTAAATTTTAAAT 3902
|||||
Db 3893 AGGCTGGACAAAGGCGCTAGTGAACCTGGCCATATGAGCTGTAAATTTTAAAT 3952
|||||
OY 3903 TTGTTTAAACATGAGCATTTTATCAAGTTAAAGGTGAGCACTTACGTGAAAAACTAT 3962
|||||
Db 3953 TTGTTTAAACATGAGCATTTTATCAAGTTAAAGGTGAGCACTTACGTGAAAAACTAT 4012
|||||
OY 3963 GTGAACATCTGTTTAACTTTCTGCTGCACTCAAAATCATTCAGCAAGTTGAGAG 4022
|||||
Db 4013 GTGAACATCTGTTTAACTTTCTGCTGCACTCAAAATCATTCAGCAAGTTGAGAG 4072
|||||
OY 4023 TCTTCAGATTTTATTAATTAAGCAACCAAGAAACATTTATGATGGAATTAATCTG 4082
|||||
Db 4073 TCTTCAGATTTTATTAATTAAGCAACCAAGAAACATTTATGATGGAATTAATCTG 4132
|||||
OY 4083 GGTGTAATTCATTTAATAATTAAGATTAATCAAGTGAATTTTGTTAATAATTTGA 4142
|||||
Db 4133 GGTGTAATTCATTTAATAATTAAGATTAATCAAGTGAATTTTGTTAATAATTTGA 4192
|||||
OY 4143 TAATTTTGTATTAATTTTATTTATTAACCTAGCTTCTCAAGATTTATTAAGAGGT 4202
|||||
Db 4193 TAATTTTGTATTAATTTTATTTATTTATTAACCTAGCTTCTCAAGATTTATTAAGAGGT 4252
|||||
OY 4203 AAAAAGTACTGAATTTTGAATTAAGTGAATTAAGTGAATTAAGTGAATTAAGTGA 4262
|||||
Db 4253 AAAAAGTACTGAATTTTGAATTAAGTGAATTAAGTGAATTAAGTGAATTAAGTGA 4312
|||||
OY 4263 AAAAA 4267
```

DB 4313 AAAA 4317
|||||
RESULT 5
ID AAN70752 standard; cDNA; 4669 BP.
XX AAN70752
AC AAN70752;
XX 21-MAY-1991 (first entry)
XX Sequence of human multi-drug resistance-1 (mdr1) cDNA from clones
DE lambda-HDR10.5 and 104.
XX
XX Chemo-therapy resistant tumour cell; P-glycoprotein; ss.
XX Homo sapiens.
OS
FH Key Location/Qualifiers
FT CDS 425..4267
FT /tag= a
XX
XX W08705943-A.
PN 08-OCT-1987.
XX
XX 26-MAR-1987; 87MO-US00758.
PF
XX 01-AUG-1986; 86US-0892575.
PR 28-MAR-1986; 86US-0845610.
XX
XX (UNIT1) UNIV OF ILLINOIS.
PI Roninson IB, Pastan IH, Gottesman MM;
DR WPI; 1987-291656/41.
DR P-PSDB; AAP70452.
XX
XX DNA for multi-drug resistance in human cells - used to detect
PT chemotherapy-resistant tumour cells and for producing
PT polypeptide(s) for diagnosis and therapy
XX
XX
PS Claim 2(a); Table 5, pp30-39; 61pp; English.
XX
XX The human multi-drug resistant KB carcinoma cell lines were used as
CC the source of the mdr1 gene nucleic acid sequences (AAN70751). To
CC obtain cDNA clones of the mdr1 gene (AAN70752), poly (A) and RNA was
CC used. Analysis of the AA sequence presented in (AAN70752) indicates
CC that the mdr1 gene product is likely to be a transmembrane protein.
CC The presence of transmembrane domains and potential glycosylation
CC sites is consistent with the mdr1 protein being related to the
CC P-glycoprotein.
XX
XX
SQ Sequence 4669 BP; 1393 A; 896 C; 1126 G; 1254 T; 0 other;
Query Match 78.1%; Score 3341.2; DB 8; Length 4669;
Best Local Similarity 87.6%; Pred. No. 0;
Matches 3758; Conservative 0; Mismatches 488; Indels 44; Gaps 8;
OY 1 GGAGGCGAGGTGCGGATGCTTGAAGAGGCGGTAAAGGAGTCAAGAGAACTT 60
DB 409 GGAGGCGAGGTGCGGATGCTTGAAGAGGCGGTAAAGGAGTCAAGAGAACTT 468
OY 61 CTGGAATAATGGCAAAAAAGTAAAAAATGAGAAAGAAAGAAACCACTGTGAG 120
DB 469 CTTTCTTAACTGAACAATAAAGTAAAGATGAAGAGAAAGAAACCACTGTGAG 528
OY 121 CACGTTTCATGTTTGCATTCATCAATGGCTGTGATGCTTATATGTTGGTGGGAC 180
DB 529 TGTATTTCATGTTTGCATTCATCAATGGCTGTGATGCTTATATGTTGGTGGGAC 588
OY 181 AATGGCTGCAATCATTCATGAGAGTCTACTCCCTCTCATGATGCTGTTTGGAAACAT 240

DB 589 TTTGGCTGCCATCATCCATGAGGCTGAGACTCTCTCATGATGCTGTGTGGAGAAAT 648
OY 241 GACGATATAGCTTTGGCAAAATGCGAGAAATTTTCAGAAACAAATTTTCAGATATATTA 300
DB 649 GACGATATATCTTTGGCAAAATGCGAGAAATTTAGAAATCTGATGT-----CAAACTATCAC 702
OY 301 TGAAGTATTTAGCAACAATATACAAACATTTTCATACACATCTGGAGAGAAATATCAC 360
DB 703 TAAATGAAGTAAATATCAATATACAGGGTCTTCAATCAATCTGAGAGAGATACACG 762
OY 361 GTATGCTATTTATTTACAGTGGATGCTGCTGCTGCTGCTGCTGCTTACATTCAGGT 420
DB 763 GTATGCTATTTATTTACAGTGGATGCTGCTGCTGCTGCTGCTGCTTACATTCAGGT 822
OY 421 TTCATCTGGAGGCTGGACAGAGAGACATATCTCAAAATTTAGAAACATTTTTCATCA 480
DB 823 TTCATTTTGGTGGCTGGACAGTGGAGACAAATATCAAAATTTAGAAACATTTTTCATCA 882
OY 481 TGTATATATGCGACAGAGAAATTTGGCTGTTGACGTGATGACGTTGGGAGCTTAACAC 540
DB 883 TGTATATATGCGACAGAGAAATTTGGCTGTTGATGATGACGATGTTGGAGCTTAACAC 942
OY 541 CCGGCTCAGACAGAGATGCTCCAAAATCAATGAAGAAATTTGGGACAAATTTGAAATGT 600
DB 943 CCGACTTACAGATGATGCTCTTAAGATTAATCAAGTATTTGTTGACAAATTTGAAATGT 1002
OY 601 CTTTCACTCAATATGCAAAATTTTTCACCGTATTATAGTGGGTTTACACGTGTTGAA 660
DB 1003 CTTTCACTCAATATGCAAAATTTTTCACGTGTTTATAGTAAATTTACACGTGTTGAA 1062
OY 661 GCTAACCTTGTATTTTGGCCATGAGCCCTGCTGAGCTTTCAGCCGATCTGGGC 720
DB 1063 GCTAACCTTGTATTTTGGCCATGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1122
OY 721 AAAGATATATCTTCTTACTGATTAAGAACTTGGCTATATGCAAAAGCTGAGCAGT 780
DB 1123 AAAGATATATCTTCTTACTGATTAAGAACTTGGCTATATGCAAAAGCTGAGCAGT 1182
OY 781 AGCTGAAGAGTGTGACAGATCAATCAAGAACTGATGCTGCTGCTGCTGCTGCTGCTGCTG 840
DB 1183 AGCTGAAGAGTGTGACAGATCAATCAAGAACTGATGCTGCTGCTGCTGCTGCTGCTGCTG 1242
OY 841 ACTTGAAGAGTGTGACAGATCAATCAAGAACTGATGCTGCTGCTGCTGCTGCTGCTGCTG 900
DB 1243 ACTTGAAGAGTGTGACAGATCAATCAAGAACTGATGCTGCTGCTGCTGCTGCTGCTGCTG 1302
OY 901 CACGCGCAACATTTCTATTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 960
DB 1303 TACAGCCAAATTTTCTATTTAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1362
OY 961 TTTCTGTATGGAGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1020
DB 1363 CTTCTGTATGGAGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1422
OY 1021 TGTCTGTATTTCTGATTAATTTGGGCTTTTATGATGAGACAGCTCCCAAGCATTTGA 1080
DB 1423 TGTATTTCTTTCTGATTAATTTGGGCTTTTATGATGAGACAGCTCCCAAGCATTTGA 1482
OY 1081 AGCATTTGCAAG 1140
DB 1483 AGCATTTGCAAG 1542
OY 1141 AAGCATTTGCAAG 1200
DB 1543 AAGCATTTGCAAG 1602
OY 1201 ATTCAAAATATGTTCACTTATGATCCCTTTCGAAAAGAGATTAAGATTTAAAGGCTCT 1260
DB 1603 ATTCAAAATATGTTCACTTATGATCCCTTTCGAAAAGAGATTAAGATTTAAAGGCTCT 1662
OY 1261 CAACCTGAAGTTCAAG 1320

Db 1663 GAACCTGAAAGTGTGACAGTGTGGGACGAGTGGCCCTGTGTGGAAACAGTGTGCGGAA 1722
Qy 1321 GAGCAGACCGCTGTGACGTGATGACAGAGGCTCTATGACCCACAGATGTGACTGTAT 1380
Db 1723 GAGCAGACAGCTGACGTGATGACAGAGGCTCTATGACCCACAGAGGGGATGTGCACTGT 1782
Qy 1381 TGATGACAGGACATTAGGACCATTAATGTAAGCAATCTGGGAAATTAATGAGTGTGCT 1440
Db 1783 TGATGACAGGACATTAGGACCATTAATGTAAGCAATCTGGGAAATTAATGAGTGTGCT 1842
Qy 1441 GAGTCAGAGACCTGTGTGTTTGTCCACCACAGATAGCTGAAGAACATTCGCTATGCGCGCA 1500
Db 1843 GAGTCAGAGACCTGTGTGTTTGTCCACCACAGATAGCTGAAGAACATTCGCTATGCGCGCA 1902
Qy 1501 AATGTCCACCATGATGATGAGAAAGCTGTTAAGSAGGCAATGCTATGATTTAT 1560
Db 1903 AATGTCCACCATGATGATGAGAAAGCTGTTAAGSAGGCAATGCTATGATTTAT 1962
Qy 1561 CATGAATACCTAATTAATTTGACACTGTGTGGAGAGAGAGGGGCCACGTGATGG 1620
Db 1963 CATGAATACCTAATTAATTTGACACTGTGTGGAGAGAGAGGGGCCACGTGATGG 2022
Qy 1621 TGAAGCAAAACAGAGATGCGCATTTGCTGGGCCCTGTGTGCAACCCCAAGATTCCT 1680
Db 2023 TGGGCAAGAGCAGAGATGCGCATTTGCTGGGCCCTGTGTGCAACCCCAAGATTCCT 2082
Qy 1681 GCTGATGAGAGCAGCTGACGCTGTGACACTGAAAGTGAAGAGTGTTCAGTGGCCCT 1740
Db 2083 GCTGATGAGAGCAGCTGACGCTGTGACACTGAAAGTGAAGAGTGTTCAGTGGCCCT 2142
Qy 1741 GGAATAGGCCAAGAAAGCGGAGTACCATTTGATGTGATGCTGATGCTGTGTGATGCTG 1800
Db 2143 GGAATAGGCCAAGAAAGCGGAGTACCATTTGATGTGATGCTGATGCTGTGTGATGCTG 2202
Qy 1801 TAAATGCCAGTGTGATGCTGCTGTTTGAATGATGAGATGATGCTGAGAAAGAAATCATGA 1860
Db 2203 TAAATGCCAGTGTGATGCTGCTGTTTGAATGATGAGATGATGCTGAGAAAGAAATCATGA 2262
Qy 1861 TGAATCATGAAAGAGAGGCAATTTACTTCAAACTGTGCACAAATGACAGCAAGAGAA 1920
Db 2263 TGAATCATGAAAGAGAGGCAATTTACTTCAAACTGTGCACAAATGACAGCAAGAGAA 2322
Qy 1921 TGAATTTAGTGTAGAAATGCGCACTGGTGAATCCAAAGTGAAGTGAATGCTTGGAAAT 1980
Db 2323 TGAATTTAGTGTAGAAATGCGCACTGGTGAATCCAAAGTGAAGTGAATGCTTGGAAAT 2382
Qy 1981 GTCTCCAAAGATTCAGAGCTCCAGTTTAATAAAAAGAGATCACTGCCAGAGATATACA 2040
Db 2383 GTCTCCAAAGATTCAGAGCTCCAGTTTAATAAAAAGAGATCACTGCCAGAGATATACA 2442
Qy 2041 TGCACCAAGGCAAGCAAGAGAAAGCTTGTACAAAAGAGAGCACTGAAATGAAATGTACC 2100
Db 2443 TGCATTCACAAAGCCCAAGACAAAGCTTGTATACAAAAGAGGCTGTGAAATGTATACC 2502
Qy 2101 TCCAGTTTCTCTGAGAGATTCGAAAGTGAACCTCACTGAAGTGGCTTATTTTGTGT 2160
Db 2503 TCCAGTTTCTCTTGTGAGAGATTAAGAGCTAAATTTAACTGAAATGCTTATTTTGTGT 2562
Qy 2161 TGTGTAATTTTGTGCTATTAATAAGAGAGCTGTGCAACAGCATTTTAAATATTTTC 2220
Db 2563 TGTGTAATTTTGTGCTATTAATAAGAGAGCTGTGCAACAGCATTTTAAATATTTTC 2622
Qy 2221 AAGGATATAGGGAATCTTTACCCGAGATGAGATCCTGTAACAAAACACAGATATAGTA 2280
Db 2623 AAGGATATAGGGAATCTTTACAAAAGATGATGATCCTGTAACAAAACACAGATATAGTA 2682
Qy 2281 CATGTTTCTGATTTGTTTCTAGCTCTTGAATTAATTTCTTTATTTATTAATTTTCTGCA 2340
Db 2683 CTGTTTCTACATTTGTTTCTAGCCCTTTGGAATTAATTTCTTTATTTATTAATTTTCTGCA 2742
Qy 2341 GGGCTTCACATTTGGCAAAAGCTGGGAGATCTCACTAAGCGGCTTCATATCATGTTT 2400
Db 2743 GGGTTTCAATTTGGCAAAAGCTGGGAGATCTCACTAAGCGGCTTCATATCATGTTT 2802

Qy 2401 CAGATCCATGCTGAGACAGATGTACGCTGTGTTGATGACCTTAATAAACACACATGGAGC 2460
Db 2803 CAGATCCATGCTGAGACAGATGTGATGTTGATGACCTTAATAAACACACATGGAGC 2862
Qy 2461 ATTGACAAACCGCTGTGCATATGATGGGCTCAAGTTAAAGGCGCTATAGGTTCCAGGCT 2520
Db 2863 ATTGACATACAGGCTCGCAATGATGTCTCAAGTTAAAGGCGCTATAGGTTCCAGGCT 2922
Qy 2521 TGTGCTATTAACCAAGATATAGCAAAATCTTGGAGACAGCATTAATTAATCTTAATCTA 2580
Db 2923 TGTGCTATTAACCAAGATATAGCAAAATCTTGGAGACAGCATTAATTAATCTTAATCTA 2982
Qy 2581 TGTGTTGGCAATTAACATTTACTGTAGCAATGTGACCATCATTTGCAATAGCAGAGAT 2640
Db 2983 TGTGTTGGCAATTAACATTTACTGTAGCAATGTGACCATCATTTGCAATAGCAGAGAT 3042
Qy 2641 TGTGTAATGAAATGTTGCTGTGACAGACAGCTGAAGATTAAGAAAGCTGAGAGAGC 2700
Db 3043 TGTGTAATGAAATGTTGCTGTGACAGACAGCTGAAGATTAAGAAAGCTGAGAGAGC 3102
Qy 2701 TGGGAAGATTTGCTACAGAGACCATGAAATCTTCGAACGTGTTTCTTTGACTGGGA 2760
Db 3103 TGGGAAGATTTGCTACAGAGACCATGAAATCTTCGAACGTGTTTCTTTGACTGGGA 3162
Qy 2761 GCGAAGTTTGAATATACATGTATGACAGAGTTTGCAGATACATACAGAAACCTTTGAG 2820
Db 3163 GCGAAGTTTGAATATGTATGTGTGACAGTTTGCAGATACATACAGAAACCTTTGAG 3222
Qy 2821 GAAAGACACATCTTGGGGGCTCATTTTCTATCACCCAGGCAATGATTTTCTCTA 2880
Db 3223 GAAAGACACATCTTGGGAATTTCTCTATCACCCAGGCAATGATTTTCTCTA 3282
Qy 2881 TGTGTGCTGTTCGCGTTTGTGTGTGCTACTTGTGTGCAATGATTCATGAACTTTCAAGA 2940
Db 3283 TGTGTGCTGTTCGCGTTTGTGTGTGCTACTTGTGTGCAATGATTCATGAACTTTCAAGA 3342
Qy 2941 TGTGTCTTGTGATTCACAGCTATTTGTCTTTGTGTGCAATGAGAGTGGGAGCTCAAGTTC 3000
Db 3343 TGTGTCTTGTGATTCACAGCTATTTGTCTTTGTGTGCAATGAGAGTGGGAGCTCAAGTTC 3402
Qy 3001 ATTGTCCTGCTATGCAAGCAAGCAAGATATGACAGAGCCACGTCATCATGATCATTTGA 3060
Db 3403 ATTGTCCTGCTATGCAAGCAAGCAAGATATGACAGAGCCACATCATGATCATTTGA 3462
Qy 3061 AAAAAGCCTCTGATGTGACAGCTACAGCCTCACGGCTCAAGCCAAATATACCTTGAAG 3120
Db 3463 AAAAAGCCTCTGATGTGACAGCTACAGCCTCACGGCTCAAGCCAAATATATGGAAG 3522
Qy 3121 AATGTGACATTTAATGAGTGTGTTCAACTATCCACTGACACAGATATCCCGCTGCT 3180
Db 3523 AATGTGACATTTGATGAGTGTGTATCAACTATCCACTGACACAGATATCCCGAGTCT 3582
Qy 3181 CCAGGGGCTGAGGCTGAGGTGAAGAAAGGCGCAGAGCGGCGCTGAGTGTGAGCTG 3240
Db 3583 TCAAGGACTGAGGCTGAGAGTGAAGAAAGGCGCAGAGCGCTGAGTGTGAGCTGAG 3642
Qy 3241 CTGTGGGAAGACACAGTTGTCAAGCTCTAAGAGGCTTCTATGACCCCTTGGCTGGTTC 3300
Db 3643 CTGTGGGAAGACACAGTGTGTCAAGCTCTAAGAGGCTTCTAAGACCCCTTGGGAGGA 3702
Qy 3301 AGTGTAAATTTGAGCAAGAGATTAAGACCTGAATGTCCAGTGGCTCCGAGACACCT 3360
Db 3703 AGTGTGTGATGAGCAAGAAATTAAGAGCTGAATGTGTAGTGGCTCCGAGACACCT 3762
Qy 3361 GGGCATGCTGTGTGAGAGGACCATCTGTTTGTGACAGCATGCGCGGAACATTTGCTTA 3420
Db 3763 GGGCATGCTGTGTGAGAGGACCATCTGTTTGTGACAGCATGCGTGTGAGAAATTTGCTTA 3822
Qy 3421 TGGAGCAAAAGCCGGGCTGATACATGAAAGAGATTAAGAGATTAAGAGAGAGAGCCAA 3480
Db 3823 TGGAGCAAAAGCCGGGCTGATACAGAGAGATCGTAGGAGAGAGAGAGAGAGCCAA 3882

Db	943	CCACACTTAAGATGATGCTCTTAAGATTAAAGAACTATTGGTGCACAAATATGGAAATGTT	1002
QY	601	CTTTCACATCAATAGCAACATTTTTCACCGGTTTATATAGTGGGTTTACACGTGGTTGAA	660
Db	1003	CTTTCAGTCAATAGCAACATTTTTCACCGGTTTATATAGTGGGTTTACACGTGGTTGAA	1062
QY	661	GCTAACCCCTTGATTTGGCCATCAGCCCTGTTCTTGGACTTTTCAGCGGCATCTGGG	720
Db	1063	GCTAACCCCTTGATTTGGCCATCAGCCCTGTTCTTGGACTTTTCAGCGGCATCTGGG	1122
QY	721	AAAGATCTATCTTCTTACTGATTAAGAACTCTTGCCCTATGCAAAAGCTGAGAGT	780
Db	1123	AAAGATCTATCTTCTTACTGATTAAGAACTCTTGCCCTATGCAAAAGCTGAGAGT	1182
QY	781	AGCTGAAGAAGCTCTAGCAGCAATCAGAACTGTGATTCCTTGGAGGACAAAGAAAGA	840
Db	1183	AGCTGAAGAAGCTCTAGCAGCAATCAGAACTGTGATTCCTTGGAGGACAAAGAAAGA	1242
QY	841	ACTGTAAAGTCAACAAATTTTAAAGAAAGCTAAAGAAATGGGATTAAGAAAGCTAT	900
Db	1243	ACTGTAAAGTCAACAAATTTTAAAGAAAGCTAAAGAAATGGGATTAAGAAAGCTAT	1302
QY	901	CACGGCCAACTTCTATATGTCGCCCTTCTTATGATCTATGCAATATGCTCTGGC	960
Db	1303	TACAGCAATATTTCTATAGGTGCTGCTTCTGCTGATCTATGCAATATGCTCTGGC	1362
QY	961	TTTCTGATGGGACCTTCTGCTCTCTCCATGCAATATTTCTATGAGCAACTACTCAC	1020
Db	1363	CTTCTGATGGGACCTTCTGCTCTCTCCATGCAATATTTCTATGAGCAACTACTCAC	1422
QY	1021	TGCTCTCTTCTGATTAATTTGGGCTTTTATGATGAGCAAGCAATCCCAACATTTGA	1080
Db	1423	TGCTCTCTTCTGATTAATTTGGGCTTTTATGATGAGCAAGCAATCCCAACATTTGA	1482
QY	1081	AGCATTTGCAACCCAGAGAGAGCTTATGAAATCTTCAAGATTAATGCAATTAACC	1140
Db	1483	AGCATTTGCAACCCAGAGAGAGCTTATGAAATCTTCAAGATTAATGCAATTAACC	1542
QY	1141	AAGCATTTGCAACCCAGAGAGAGCTTATGAAATCTTCAAGATTAATGCAATTAACC	1200
Db	1543	AAGCATTTGCAACCCAGAGAGAGCTTATGAAATCTTCAAGATTAATGCAATTAACC	1602
QY	1201	ATTCAAAAATGTTCACTTCACTTACCTTCTGCAAAAAGTATAGATCTTAAGAGCTT	1260
Db	1603	ATTCAAAAATGTTCACTTCACTTACCTTCTGCAAAAAGTATAGATCTTAAGAGCTT	1662
QY	1261	CAACCTGAAGTTTCAGAGTGGGACAGAGTGGGCTGTTGGGAAACAGTGGCTCGGAA	1320
Db	1663	CAACCTGAAGTTTCAGAGTGGGACAGAGTGGGCTGTTGGGAAACAGTGGCTCGGAA	1722
QY	1321	GAGCAGACCGTGCAGTGCAGAGGCTCTATGACCCACAGATGGCATGAGTCTGTAT	1380
Db	1723	GAGCAGACCGTGCAGTGCAGAGGCTCTATGACCCACAGATGGCATGAGTCTGTAT	1782
QY	1381	TGATGGACAGCAATTTAGCAACATTAATTTAAGGCACTCTCGGGAATTTACTGGTGGT	1440
Db	1783	TGATGGACAGCAATTTAGCAACATTAATTTAAGGCACTCTCGGGAATTTACTGGTGGT	1842
QY	1441	GAGTCAGAAGCTGTGTTGTTGCCACACAGATAGCTGAAAACATTCGCTATGGCCGCA	1500
Db	1843	GAGTCAGAAGCTGTGTTGTTGCCACACAGATAGCTGAAAACATTCGCTATGGCCGCA	1902
QY	1501	AAATGTCACATGATGATGATGAAAGCTTTAAGAAAGCTATGGCATTTGATTTAT	1560
Db	1903	AAATGTCACATGATGATGATGAAAGCTTTAAGAAAGCTATGGCATTTGATTTAT	1962
QY	1561	CATGAAGTACTATTAATTTGACACTCTGTTGGAGAGAGGGCCACGCTAGCTG	1620
Db	1963	CATGAAGTACTATTAATTTGACACTCTGTTGGAGAGAGGGCCACGCTAGCTG	2022
QY	1621	TGGACAGAAACAGAAATCGCATTTGCTGGGCGCTGTTGCCAACCCAGATTTCTTCT	1680

Db	2023	TGGGACGAAGCAGAGAGATGCCATTTGCAGTGCCTCGCTGGTGCACAACTCCATCTCCT	2082
QY	1681	GCTGGATGAGGCAAGCTGACCTCTGACACTGAAAGTGAAGCAGTGGTTCAGTGGCCT	1740
Db	2083	GCTGGATGAGGCAAGCTGACCTCTGACACTGAAAGTGAAGCAGTGGTTCAGTGGCCT	2142
QY	1741	GGATTAAGGCAAGAAAGGCGACTATCATTTGATGATGATGATGATGATGATGATGATG	1800
Db	2143	GGATTAAGGCAAGAAAGGCGACTATCATTTGATGATGATGATGATGATGATGATGATG	2202
QY	1801	TATGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG	1860
Db	2203	TATGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG	2262
QY	1861	TGACATCATGAAGAGAGGCAATTTACTTCAAACTTGTCAACATGACAGACAGAGAAA	1920
Db	2263	TGACATCATGAAGAGAGGCAATTTACTTCAAACTTGTCAACATGACAGACAGAGAAA	2322
QY	1921	TGAAATTTGATTAAGAAATGTCACCTGATGATGATGATGATGATGATGATGATGATG	1980
Db	2323	TGAAATTTGATTAAGAAATGTCACCTGATGATGATGATGATGATGATGATGATGATG	2382
QY	1981	GCTCCAAAAGTTTCAAGGCTCAGTTTATTAAGAAAGATCACTGCAAGAGATATCA	2040
Db	2383	GCTCCAAAAGTTTCAAGGCTCAGTTTATTAAGAAAGATCACTGCAAGAGATATCA	2442
QY	2041	TGCACACAGAGCCAGACAGAAAGCTTGGTACAAAAGAGAGCTTGAATGGAATGTACC	2100
Db	2443	TGCACACAGAGCCAGACAGAAAGCTTGGTACAAAAGAGAGCTTGAATGGAATGTACC	2502
QY	2101	TCCAGTTCTCTTGGAGGATGTCGAACTGCACTCACTGATGATGATGATGATGATGATG	2160
Db	2503	TCCAGTTCTCTTGGAGGATGTCGAACTGCACTCACTGATGATGATGATGATGATGATG	2562
QY	2161	TGCTATATTTTGTGCTATTAATTAAGAGGCTCGACACAGCATTTTCAATATTTTC	2220
Db	2563	TGCTATATTTTGTGCTATTAATTAAGAGGCTCGACACAGCATTTTCAATATTTTC	2622
QY	2221	AAGGATTAATAGGATCTTTACCGAGATGAGATCTCTGAAACAAACAAACAGATATGTA	2280
Db	2623	AAGGATTAATAGGATCTTTACCGAGATGAGATCTCTGAAACAAACAAACAGATATGTA	2682
QY	2281	CATGTTTCTGATTTGTTTCTGACCTGGAATTTCTTTTATTTATTTTCTTCTTCTTCT	2340
Db	2683	CATGTTTCTGATTTGTTTCTGACCTGGAATTTCTTTTATTTTATTTTCTTCTTCTTCT	2742
QY	2341	GGGCTTACATTTGGCAAGCTGGGAGATCTCACTAACCGCTTGGATCATGATGTTT	2400
Db	2743	GGGCTTACATTTGGCAAGCTGGGAGATCTCACTAACCGCTTGGATCATGATGTTT	2802
QY	2401	CAGATCATGCTGAGACAGATGTCAGCTGTTGATGACCTTAATAAACACACACTGGAGC	2460
Db	2803	CAGATCATGCTGAGACAGATGTCAGCTGTTGATGACCTTAATAAACACACACTGGAGC	2862
QY	2461	ATTGACACAGGCTTGGCAATGATGCGCTCAAGTTAAAGGCGCTTAAGGCTTCAAGGCT	2520
Db	2863	ATTGACACAGGCTTGGCAATGATGCGCTCAAGTTAAAGGCGCTTAAGGCTTCAAGGCT	2922
QY	2521	TGCTCATTTACCGAATATAGCAAACTTGGGACAGGCAATTAATATCTTAATCTA	2580
Db	2923	TGCTCATTTACCGAATATAGCAAACTTGGGACAGGCAATTAATATCTTAATCTA	2982
QY	2581	TGGTTGGCAATTAACACTTTACTTACCAATTTGACCAATTTGCAATGATGAGAGAGT	2640
Db	2983	TGGTTGGCAATTAACACTTTACTTACCAATTTGACCAATTTGCAATGATGAGAGAGT	3042
QY	2641	TGTTGAATGAATGTTGTTGTCGACAGACCTGAAAGTAAAGAAAGCTAGAAAGAGC	2700
Db	3043	TGTTGAATGAATGTTGTTGTCGACAGACCTGAAAGTAAAGAAAGCTAGAAAGAGC	3102
QY	2701	TGGAGAGATTTGCTACAGAGCCTGAAACTTCCGAATGTTGTTCTTCTTCTTCTTCTTCT	2760
Db	3103	TGGAGAGATTTGCTACAGAGCCTGAAACTTCCGAATGTTGTTCTTCTTCTTCTTCTTCT	3162

OY	2761	GCACAGTTTGAAATACATGTATGACACAGAGTTTGCAAGTACCATACAGAAACTGTTGAG	2820
Db	3163	GCAGAGTTTGGACATATGTATGCTCAGAGTTTGACAGTACCATACAGAAACTCTTTGAG	3222
OY	2821	GAAGACACACATCTTGGGGCTCTCATTTCTTCTATACACCCAGGCAATGATATATTTTCCTA	2880
Db	3223	GAAGACACACATCTTGGGAATTAATTTTCTTCCACCCAGGCAATGATATATTTTCCCTA	3282
OY	2881	TGCTGCGTGTTCGGGTTTGGTGTGCTTACTTGTGTGGCAATGAGTTCATGACCTTTTGACA	2940
Db	3283	TGCTGAGATGTTTCCGGTTTGGAGCCCTACTTGTGTGGCACATTAATATATAGCTTTTGAGA	3342
OY	2941	TGCTCTTTTGGATATCTTCAGCTATTTGCTTTGGTGCCTATGGCAGTGGGGCAGTTC	3000
Db	3343	TGTTTGTTAGTATTTTACGCTGTTCCTTTGGTGGCATGCGGTGGGGCAATCAGTTC	3402
OY	3001	ATTGTGCTCCGACTATGCCAAAGCCAAAGTATCAGAGCCCAAGTATCTATGATCAATTGA	3060
Db	3403	ATTGTGCTCCGACTATGCCAAAGCCAAATATCAGAGCCCAATCATATGATCAATTGA	3462
OY	3061	AAAAACCCCTGATGAGACAGCTACAGCCCTCACGGCTCAGCCCAATATCTTTGGAAG	3120
Db	3463	AAAAACCCCTTATATTGACAGCTACAGCAGGAAGGCCCTATATGCCAAACATTTGGAAG	3522
OY	3121	AAATGTGACATTTTAAATGAGTGTGTCAACATATCCACTGACACAGCATCCCCGCT	3180
Db	3523	AAATGTACATTTTGGTGAAGTGTATTAACATATCCACCCGACCCGACATCCAGTGT	3582
OY	3181	CCATGGGCTAGGCTGAGTGTGAAGAAAGGGCCAGAGCGTGGCCCTCTGATAGTATACATGG	3240
Db	3583	TCAGGACATAGGCTGAGGTGAAGAAAGGGCCAGAGCGTGGCTGTGGTGGGACAGATGG	3642
OY	3241	CTGTGGGAAGAGCACATTTGTACGCTCCAGAGCGCTTCTATGACCCCTTGGCTGTTTC	3300
Db	3643	CTGTGGGAAGAGCACATGTTGTACGCTCCAGAGCGCTTCTACGACCCCTTGGGCAAGGAA	3702
OY	3301	AGTGTATATTTATGGCAAAAGATAAAGCACCTGAAATGTCCAGTGGCTCCGAGCACACT	3360
Db	3703	AGTGTCTTATATGGCAAAAGAAATAAAGCCAGTAATGTTCAGTGGCTCCGAGCACACT	3762
OY	3361	GGGACATGCTGTCTGAGAGCCCATCTCGTTTGAATCTGACGATTTGCCGAGAACATTTGCCTA	3420
Db	3763	GGGACATGCTGTCTGAGAGCCCATCTCGTTTGAATCTGAGAACATTTGCCCTA	3822
OY	3421	TGAGAGCAACAGCCGGGTCTGTATCAACATGAAGAGATTATCAGAGCCAGCCAAAGGSCCA	3480
Db	3823	TGAGAGCAACAGCCGGGTGTGTCAACAGAAATATGTAGGGCAGCAAAAGAGGSCCA	3882
OY	3481	CATACACCACTTATCTGAGACACTCCCTGAGAAATACACACAGAGTATAGAGACAAAG	3540
Db	3883	CATACATGGCTTATCATGATCACTCCTCAATTAATATAGCACTTAATATAGAGACAAAG	3942
OY	3541	AACCCAGCTCTCTGTGTGGCCAGAAACAGCGATTGSCATATGCTGCGCTCTTTAGACA	3600
Db	3943	AACCTCAGCTCTGTGTGGCCAGAAACAGCATTTCCATCTGTGCTGCTCTTTTAGACA	4002
OY	3601	GCCCATATTTTCTTTTGGATGAAGCTACATCACTCTGGATTCAGAAATGAAAAGT	3660
Db	4003	GCCCATATTTTCTTTTGGATGAAGCCACGTCACCTCTGGATTCAGAAATGAAAAGT	4062
OY	3661	TGTCCAAAGAGCCCTGACAAACCCAGAAAGGCGCACCTGATTTGATGCCACCG	3720
Db	4063	TGTCCAAAGAGCCCTGACAAACCCAGAGAGGCGCACCTGATTTGATTTGCCACCG	4122
OY	3721	CTTGTCCACCAATCCAGATCCAGATTTTAATAGTGTGTTTCAGATGGCAAAAGTCAGGA	3780
Db	4123	CTTGTCCACCAATCCAGATCCAGATTTAATAGTGTGTTTCAGATGGGAGAGTCAAGGA	4182
OY	3781	GCATGGGACCAATCAAGAGTGTGTGGCCCAAGAAAGCATATTTTTCATGTTCAGTGT	3840
Db	4183	GCATGGGACCAATCAAGAGTGTGTGGGACAGAAAGGACATTTTTCATGTTCAGTGT	4242

QY	3641	CCAGCGTCGAGCAAGAAACGCGTAGTGAACCTGTGGCCATATGAGCTGTTAAATATTTTAAAT	3900
Db	4243	CCAGGCGTGGAAACAAAGCCGCACTGAACCTGATGTAGATGTTAAATACCTTTAAAT	4302
QY	3901	ATTGTGTTAAAAACATGCGATTTAAATCAAAAGTTAAAGTGAGCAGCTTACGTGAAAACT	3960
Db	4303	ATT--TGTAGATATGACATTTATTCAAAGTTAA-----AGCAACACTTACAGAAAT	4355
QY	3961	ATGTAGAACTACCGTGTGTTAACTTCTTGCTGCACACGTGAGATCATTCACCAAGTTGAG	4020
Db	4356	ATGAAGAGGATCTGTTTAACATTTCTCAGTCAGTTCAGAGCTTC-----	4403
QY	4021	AGCTTCAGATTTTATTAATTAAGAGACCA--AAGAAACATTAATCGATGAAATAAAT	4078
Db	4404	-----AGAGACTTCGTATTAAAGGAACAGATGAGAGACATCATCAAGTGAGAGAAAT	4458
QY	4079	ACTGCTTATTTGTCATTAATAAATTATAGAGTAATTCAAAGTACATTTTGTATTAAT	4138
Db	4459	CATAGCTTTAAACGTCATTTATTAATTTATATACGAATTTAAAGTAAATTTTAAAGATPAA	4518
QY	4139	TGTATATATTTTGGTTATATTTT-----ATTGTACTTACAGCTTGTGCTGAAGAT	4190
Db	4519	ATGCTGAATTTGGTTATATTTTCCACTTGGACTGTAACTGACTGCTTGTCTAAAGAT	4578
QY	4191	TATAGAAAGGTGATTAATAAGTACTG--AATGTTGAAATAAAGTCTAGCTATTAATAAAGTAA	4249
Db	4579	TATAGAAAGTAGAATAAAGTATGAAATGTTGCATAAAGT--TCTATTAATAAAGCTPAA	4635
QY	4250	ACTTTATATCAAAAAAAGAAAAA 4279	
Db	4636	ACTTTCATGTGAAAAAAGAAAAA 4665	
RESULT 8			
AAQ04522			
ID	AAQ04522	standard; DNA; 4378 BP.	
XX	AAQ04522:		
AC			
XX	01-OCF-1990 (first entry)		
DI			
XX	MultiDrug Resistance A1 gene.		
DE			
XX	MultiDrug Resistance A1 gene; drug resistance of cancer cells; ss.		
KW			
XX	Homo sapiens.		
OS			
XX	Location/Qualifiers		
EH	Key	138..3981	
FT	CDS	/**tag= a	
XX			
XX	JP02100680-A.		
XX			
PD	12-APR-1990.		
XX			
PF	05-OCT-1988; 88JP-0251475.		
XX			
PR	05-OCT-1988; 88UO-0251475.		
XX			
PA	(SUNR) SUNTORY LTD.		
XX			
DR	WPI: 1990-159707/21.		
XX	P-PSDB: AAR04868.		
XX			
PT	MDR related gene derived from human normal cells -		
XX	resistance of cancer cells		
XX	resistance of cancer cells		
XX	Disclosure; : P: Japanese.		
XX			
CC	The gene is useful for diagnosis of drug resistance of cancer		
CC	cells.		
XX			

Sequence 4378 BP; 1313 A; 819 C; 1058 G; 1188 T; 0 other;

Query Match 77.9%; Score 3334.8; DB 11; Length 4378;
Best Local Similarity 87.5%; Pred. No. 0;
Matches 3754; Conservative 0; Mismatches 492; Indels 44; Gaps 8;

QY 1 GGAGCGCAGAGTGGGATGATCTCTGAAGAGAGGCCGTAAGGGAGGCGAGAGAAAGACTT 60
DB 122 GGAGCGCAGAGTGGGATGATCTCTGAAGAGAGGCCGTAAGGGAGGCGAGAGAAAGAA 181
QY 61 CTGAAAATGGGCAAAAAAGTAAAAAATGAGAAAGAAAGAAAGAAAGAAAGAAAGAA 181
DB 182 CTTTAAAACTGAAACATTAAGTAAAGAAAGTAAAGAAAGAAAGAAAGAAAGAAAGAA 241
QY 121 CACGTTTGCAATGTTGGCTATTCAAATGGCTGTAGTGTGTATAGTTGGTGGGAC 180
DB 242 TGTATTTCAATGTTGGCTATTCAAATGGCTGTAGTGTGTATAGTTGGTGGGAC 301
QY 181 AATGGCTGCATCATCATGAGCTGACCTCCCTCATGATGCTGTTTGGAAACAT 240
DB 302 TTTGGCTGCATCATCATGAGCTGAGCTTCTCTCATGATGCTGTTGGAGAAAT 361
QY 241 GACAGATAGCTTGCAGAAATGCAAGAAATTCAGAAAGAAAGAAAGTTCAGTTAATTA 300
DB 362 GACAGATAGCTTGCAGAAATGCAAGAAATTCAGAAAGTTCAGTTAATTA 415
QY 301 TGAAGATTTAGCAACAATACACAACATTTCTATCAACCATGAGAGAGAAATGACAC 360
DB 416 TAATAGAGTATATCAAAATGATACAGAGGTTCTCATGAATCTGAGAGAAACATGAC 475
QY 361 GTATGCTATTTATTAAGTGGGATGCTGTGGGCTGCTGGTGGCTCTTAATCAAGT 420
DB 476 GTATGCTATTTATTAAGTGGGATGCTGTGGGCTGCTGGTGGCTCTTAATCAAGT 535
QY 421 TTGATTTGGTGGCTGGGAGAGAGACAGATCTCAAAATTAAGAAACATTTTGA 480
DB 536 TTGATTTGGTGGCTGGGAGAGAGACAGATCTCAAAATTAAGAAACATTTTGA 595
QY 481 TGCATATGCGACAGAGATTTGGCTGTTTGAAGTTCATGACGTTGGGAGCTTAAC 540
DB 596 TGCATATGCGACAGAGATTTGGCTGTTTGAAGTTCATGACGTTGGGAGCTTAAC 655
QY 541 CGGCTCACAGAGATGCTCAAAATCAATGAAGAAATTTGGACAAATTTGGATGT 600
DB 656 CGGCTCACAGAGATGCTCAAAATCAATGAAGAAATTTGGACAAATTTGGATGT 715
QY 601 CTTTCACTCAATAGCAACATTTTTCACCGGTTTATAGTGGGTTTACAGTGGTTGAA 660
DB 716 CTTTCACTCAATAGCAACATTTTTCACCGGTTTATAGTGGGTTTACAGTGGTTGAA 775
QY 661 GCTAACCCCTGTATTTGGGCAATCAAGCCCTGTCTTGGACTTCAAGCCGACCTGGGC 720
DB 776 GCTAACCCCTGTATTTGGGCAATCAAGCCCTGTCTTGGACTTCAAGCCGACCTGGGC 835
QY 721 AAAGATCTATCTTCAATTTACTGATTAAGAACTTTGGCTATGCAAGAAAGTGGAGAGT 780
DB 836 AAAGATCTATCTTCAATTTACTGATTAAGAACTTTGGCTATGCAAGAAAGTGGAGAGT 895
QY 781 AGCTGAAGAGTGTTCAGCAATCAAGAACTGTGATTCCTTTGGAGAGCAAAAGAAAGA 840
DB 896 AGCTGAAGAGTGTTCAGCAATCAAGAACTGTGATTCCTTTGGAGAGCAAAAGAAAGA 955
QY 841 ACTTGAAGAGTACAAACAAAATTTAGAGAGCTTAAGGAATTTGGATTAAGAAAGCTAT 900
DB 956 ACTTGAAGAGTACAAACAAAATTTAGAGAGCTTAAGGAATTTGGATTAAGAAAGCTAT 1015
QY 901 CAGCGCAACATTTTATTTGGCTGCTTCTTATGATCATGATCATGATCTGGC 960
DB 1016 TACAGCAATATTTCTATAGTCTGCTTCTGCTGATCATGATCATGATCTGGC 1075
QY 961 TTTCTGATGGAGCTCTGCTGCTCTCAAGTAATATTTGATGGAAGTCAAGTCACTAC 1020
DB 1076 CTTTGTGATGGAGCACTTGGTCTCTCAAGTAATATTTGATGGAAGTCAAGTCACTAC 1135

QY 1021 TGTCTTCTTTCTGATTAATTTGGGGCTTTTATGTTGGACAGGCAATCCCAAGCATTTGA 1080
DB 1136 TGTATTTCTTTCTGATTAATTTGGGGCTTTTATGTTGGACAGGCAATCCCAAGCATTTGA 1195
QY 1081 AGCATTTGCAACGCAAGAGGAGCAAGCTTATGAATCTCAAGATTAATGACATAAAC 1140
DB 1196 AGCATTTGCAACGCAAGAGGAGCAAGCTTATGAATCTCAAGATTAATGATTAAGCC 1255
QY 1141 AAGCATTTGCAAGCTTATGGAAGAGTGCATTAACAGATTAATTAAGGAATTTGGA 1200
DB 1256 AAGTATTTGCAAGCTTATGGAAGAGTGCATTAACAGATTAATTAAGGAATTTGGA 1315
QY 1201 ATTCAAAAATGTTCACTTCAATTAACCTTCTCGAAAAAGATTAAGATTTAAAGTCT 1260
DB 1316 ATTCAAAAATGTTCACTTCAATTAACCTTCTCGAAAAAGATTAAGATTTAAAGTCT 1375
QY 1261 CAACCTGAAGTTCAGAGTGGGAGAGAGTGGGCTGTTGGGAGACAGTGGCTGGGAA 1320
DB 1376 GAACCTGAAGTTCAGAGTGGGAGAGAGTGGGCTGTTGGGAGACAGTGGGAA 1435
QY 1321 GAGCAGACCGGTGAGAGTGCAGAGAGCTCTATGACCCACAGATAGGATGCTGTAT 1380
DB 1436 GAGCAGACCGGTGAGAGTGCAGAGAGCTCTATGACCCACAGAGATAGGATGCTGTAT 1495
QY 1381 TGATGACAGAGACATTTAGGACCATTAATGTAAGGCATCTTGGGAAATTAATCTGTGCT 1440
DB 1496 TGATGACAGAGACATTTAGGACCATTAATGTAAGGTTCTTACGGGAAATCAATGTTGCT 1555
QY 1441 GAGCAGAGACCTGTGTGTTTCCACACAGATAGCTGAAAAACATTCCTTATGAGCCGGA 1500
DB 1556 GAGCAGAGACCTGTGTGTTTCCACACAGATAGCTGAAAAACATTCCTTATGAGCCGGA 1615
QY 1501 AAATGTCACCATGATGATGATGGAAGAGCTGTAAAGAACCAATGCTATGATTTAT 1560
DB 1616 AAATGTCACCATGATGATGATGGAAGAGCTGTAAAGAACCAATGCTATGATTTAT 1675
QY 1561 CATGAATACCTCAATTAATTTGACACTGCTGTGGAGAGAGAGGAGCCAGAGTGG 1620
DB 1676 CATGAATACCTCAATTAATTTGACACTGCTGTGGAGAGAGAGGAGCCAGAGTGG 1735
QY 1621 TGACAGAAACAGAGAAATGCGCATGCTGGGAGCTGTGGCAACCCCAAGATTTCTGT 1680
DB 1736 TGAGGAGAGAGAGAGATGCGCATGCTGGGAGCTGTGGCAACCCCAAGATTTCTGT 1795
QY 1681 GCTGAGTGAAGGACAGTACGCTTGACACTGAAAGTGAAGCAGTGGTACAGTGGCCCT 1740
DB 1796 GCTGAGTGAAGGACAGTACGCTTGACACTGAAAGTGAAGCAGTGGTACAGTGGCCCT 1855
QY 1741 GGATGAAGGCCAGAAAGAGCCGAGCTACCATTTGATAGCTCATGCTTTGTACAGTTGG 1800
DB 1856 GGATGAAGGCCAGAAAGAGTTCGACACCATTTGATAGCTCATGCTTTGTACAGTTGG 1915
QY 1801 TAAAGCCATGCTCATGCTGTGTTTGAATGATGAGTCAATTTGGAGAAAGAAATCATGA 1860
DB 1916 TAAAGCCATGCTCATGCTGTGTTTGAATGATGAGTCAATTTGGAGAAAGAAATCATGA 1975
QY 1861 TGAACATGAAAGAGAGGCAATTTACTTAACCTTGTCAAAATGCAAGCAAGAGAAA 1920
DB 1976 TGAACATGAAAGAGAGGCAATTTACTTAACCTTGTCAAAATGCAAGCAAGAGAAA 2035
QY 1921 TGAATTTGAGTTGAAAATGCCACTGTGAATCCAAAAGTGAAGTGAATGATGCTTGGAAAT 1980
DB 2036 TGAATTTGAGTTGAAAATGCCACTGTGAATCCAAAAGTGAAGTGAATGATGCTTGGAAAT 2095
QY 1981 GCTTCCAAAAGATTTGAGGTCAGTTTAAATAAAGAAAGATCAACTGCAGAGATATACA 2040
DB 2096 GCTTCCAAAAGATTTGAGGTCAGTTTAAATAAAGAAAGATCAACTGCAGAGATATACA 2155
QY 2041 TGACACCAAGGCCAAGACAGAAAGCTGTGTAACAAAGAGAGTGTGAATGAGATATACC 2100
DB 2156 TGATATCAAGGCCAAGACAGAAAGCTGTGTAACCAAGAGAGTGTGTGAATGAGATATACC 2215

```
OY 2101 TCCAGTTTCCTCTGAGAGATTCGAAAGCTGAACTCACTGAATGAGCCCTTATTTGCTG 2160
|||||
DB 2216 TCCAGTTTCCTCTGAGAGATTCGAAAGCTCAAAATTTACGAAATGGCTTATTTGCTG 2275
|||||
OY 2161 TGGTATATTTTGGCTATTTATTAACGAGGCTTCGAACCAAGATTTTCATATATTTTC 2220
|||||
DB 2276 TGGTATATTTTGGCTATTTATTAATGAGGCTTCGAACCAAGATTTTCATATATTTTC 2335
|||||
OY 2221 AAGGATTAATGAGGATCTTACCCGAGATCCCTGAACAAACGAGCAAGTAATGA 2280
|||||
DB 2336 AAGGATTAATGAGGATTTTACAGAAATGATGATCCGAAACAAACGAGCAAGTAATGA 2395
|||||
OY 2281 CATGTTTCTGATATGTTTCTAGTCCCTGGAATTAATTTCTTTATTTACATTTTTCCTCA 2340
|||||
DB 2396 CTGTTTTCACATTTGTTCTAGCTTCGAAATTTATTTCTTTATTTACATTTTTCCTCA 2455
|||||
OY 2341 GGGCTTCACATTTGGCAAAAGCTGGGAGATCCCTCAAGGCGCTCGATACATGTTTTT 2400
|||||
DB 2456 GGGTTTCACATTTGGCAAAAGCTGGAGATCCCTCAACCAAGGCGCTCCGATACATGTTTTT 2515
|||||
OY 2401 CAGATTCATGCTGTGAGACAGAGATGTCAAGTGTGATGACCTTAAACACCACTGGAGC 2460
|||||
DB 2516 CCGATTCATGCTGTGAGACAGAGATGTGAGTGTGATGACCTTAAACACCACTGGAGC 2575
|||||
OY 2461 ATTGACAAACAGGCTTGGCAATGATGAGGCTCAAGTTAAAGGCGCTATAGTTCCAGCT 2520
|||||
DB 2576 ATTGACAAACAGGCTTGGCAATGATGAGGCTCAAGTTAAAGGCGCTATAGTTCCAGCT 2635
|||||
OY 2521 TGTCTGCAATTTACCCAGAAATATAGCAAAATCTTGGGACAGGATATATATTCCTTATCTA 2580
|||||
DB 2636 TGTCTGCAATTTACCCAGAAATATAGCAAAATCTTGGGACAGGATATATATTCCTTATCTA 2695
|||||
OY 2581 TGTCTGCAATTTACACATTTTACTTGTAGCAATTTGATACCATCATCTGCAATGACAGAGT 2640
|||||
DB 2696 TGTCTGCAATTTACACATTTTACTTGTAGCAATTTGATACCATCATCTGCAATGACAGAGT 2755
|||||
OY 2641 TGTCTGCAATTTAGAAATGTTGTCTGGAACAAGCTGAAGATTAAGAGGTAGAAAGAGC 2700
|||||
DB 2756 TGTCTGCAATTTAGAAATGTTGTCTGGAACAAGCTGAAGATTAAGAGGTAGAAAGAGC 2815
|||||
OY 2701 TGGGAAGATTTGCTACAGAAAGCATGGAACCTCCGAATGTTTCTTGTGACTGGGA 2760
|||||
DB 2816 TGGGAAGATTTGCTACAGAAAGCATGGAACCTCCGAATGTTTCTTGTGACTGGGA 2875
|||||
OY 2761 GCAGAGTTTGAATACATATATGACAAGATTTGCAAGTACCATACAGAAAGCTTTGAG 2820
|||||
DB 2876 GCAGAGTTTGAATACATATATGATGATGAGTTTGCAGAGTTCACATACAGAAAGCTTTGAG 2935
|||||
OY 2821 GAAAGCACATCTTGGGAGGCTCATTTTCTATCACCCAGGCAATGATATTTTCTTA 2880
|||||
DB 2936 GAAAGCACATCTTGGGAGGCTCATTTTCTATCACCCAGGCAATGATATTTTCTTA 2995
|||||
OY 2881 TGTGAGCTGTTTCCGGTTTGGTGGCTTACTTGTGGAATGAGTTGCAAGATTTTCAGA 2940
|||||
DB 2996 TGTGAGAGTGTTCGGGTTTGGAGCTTACTTGTGGAATGAGTTGCAAGATTTTCAGA 3055
|||||
OY 2941 TGTCTTTTGGTATTTCACTATTTGCTTGTGATGAGGAGTGGGCAAGTCAAGTTTC 3000
|||||
DB 3056 TGTCTTTTGGTATTTCACTATTTGCTTGTGATGAGGAGTGGGCAAGTCAAGTTTC 3115
|||||
OY 3001 ATTTGCTCTGACATGACCAAGCAAGATATGACAGCCCAAGCTACATCATGATGATTTGA 3060
|||||
DB 3116 ATTTGCTCTGACATGACCAAGCAAGATATGACAGCCCAAGCTACATCATGATGATTTGA 3175
|||||
OY 3061 AAAAGCCCTCTGATGACAGCTACAGCCCTCAAGGCGCTCAAGCCAAATACGTTGGAAG 3120
|||||
DB 3176 AAAAGCCCTCTGATGACAGCTACAGCCCAAGGAGGCTTATGCGAACAACATTTGGAAG 3235
|||||
OY 3121 AAATGCAATTTATGAGGCTGTGTTCAACTATCCACTGACACAGACATCCCGTGTG 3180
|||||
DB 3236 AAATGCAATTTGTTGTAAGTTGTAATTCACATATCCACCGACGAGATCCCACTGCT 3295
|||||
OY 3181 CCAGGAGCTGAGCTGAGGTGAAGAGGCGACAGCCCTGACCTGTAGTAGCAGTGG 3240
|||||

DB 3296 TCAGGAGCTGAGCTGAGAGTGAAGAGGCGCACAGACGCTGGCTGTGGGACAGACTG 3355
|||||
OY 3241 CTGTTGGAGAGACACAGTTGTTTACGCTCTAGAGCGCTTCTATATACCCCTTGGCTGTT 3300
|||||
DB 3356 CTGTTGGAGAGACACAGTGTGTCCAGCTCTGTGAGGCGGTTCTACAGACCCCTTGGAGGAA 3415
|||||
OY 3301 AGTGTATTTGATGCAAAAGATTAAGACACTGAATGTCCAGTGGGCTCCGACACCT 3360
|||||
DB 3416 AGTGTCTCTGATGCAAAAGATTAAGAAATTAAGGACTGAATGTCTGAGTGGCTCCGACACCT 3475
|||||
OY 3361 GGGGATGCTGTCCAGAGAGCCCATCTGTTTGAATGACAGCATTTGCCGGAACATTTGCTA 3420
|||||
DB 3476 GGGGATGCTGTCCAGAGAGCCCATCTGTTTGAATGACAGCATTTGCCGGAACATTTGCTA 3535
|||||
OY 3421 TGGAGACACAGCCGCTGTATACATGAAGATTTATGACAGGAGCCCAAGAGCCCA 3480
|||||
DB 3536 TGGAGACACAGCCGCTGTATACATGAAGATTTATGACAGGAGCCCAAGAGGAGCCCA 3595
|||||
OY 3481 CATACACACTTCATGAGACACTCCCTGAGAAATACAAACCAAGATGAGACAAAG 3540
|||||
DB 3596 CATACACTCTTTCATGAGACTGCTATTAATATAGCACTAAAGTGAAGAGCAAAAG 3655
|||||
OY 3541 AACCAAGCTCTGTGTGCGCAAGAAACAGCCATTTGCCATGCTGCGCTTGTAGACA 3600
|||||
DB 3656 AACCAAGCTCTGTGTGCGCAAGAAACAGATTTGCCATGCTGCTGTTAGACA 3715
|||||
OY 3601 GCGTCAATTTTGTGTTGATGAGAGTACATCAGCTCTGTGATACGAAAGTCAAAAGT 3660
|||||
DB 3716 GCGTCAATTTTGTGTTGATGAGAGTACATCAGCTCTGTGATACGAAAGTCAAAAGT 3775
|||||
OY 3661 TGTCCAAGAGCCCTGTGAGCAAAAGCCAGAGAGGCGGACCTGATGATGATCCGACCG 3720
|||||
DB 3776 TGTCCAAGAGCCCTGTGAGCAAAAGCCAGAGAGGCGGACCTGATGATGATCCGACCG 3835
|||||
OY 3721 CTGTCCACATTCAGAAATGCAATTTAATAGTGTGTTTCAAGATGCAAAAGTCAAAAG 3780
|||||
DB 3836 CTGTCCACATTCAGAAATGCAATTTAATAGTGTGTTTCAAGATGCAAAAGTCAAAAG 3895
|||||
OY 3781 GCATGACACATCAACAGCTGCTGGCCCAAGAAAGCATATTTTCCATGCTGACTGT 3840
|||||
DB 3896 GCATGACACATCAACAGCTGCTGGCCCAAGAAAGCATATTTTCCATGCTGACTGT 3955
|||||
OY 3841 CCAGGCTGAGCAAGGCGTATGAACTGTGGCATATGAGCTGTTAAATTTTAAAT 3900
|||||
DB 3956 CCAGGCTGAGCAAGGCGCAAGGCAAGTGTGACTGTATGAGTGTAAATTTTAAAT 4015
|||||
OY 3901 ATTTGTGTTAAACATGCACTTTTAAATCAAAAGTTAAAGGTGACCTTACTGCAAAACT 3960
|||||
DB 4016 ATTTGTGTTAAACATGCACTTTTAAATCAAAAGTTTAAAGGTGACCTTACTGCAAAACT 4068
|||||
OY 3961 ATGTAAGAACTGCTGTTTAACTTTTCTGCTGCACTGAAATCAATTCACCAAGTTGAG 4020
|||||
DB 4069 ATGTAAGAACTGCTGTTTAACTTTTCTGCTGCACTGAAATCAATTCACCAAGTTGAG 4116
|||||
OY 4021 AGCTTCGCAATTTATTAATTAAGAACCAAA--AGCAACATTAATTCATGAGTAAT 4078
|||||
DB 4117 AGCTTCGCAATTTATTAATTAAGAACCAAGAGTGAAGACATCAATCAAGTGAAGAAAT 4171
|||||
OY 4079 ACTGCTGTTAATGATTAATTAATTAATTAAGAGTGAAGTGAAGTGAAGTGAAGTGA 4138
|||||
DB 4172 CGTAGTTTAACTGATTAATTAATTAATTAAGAAATTAAGTGAAGTGAAGTGAAGTGA 4231
|||||
OY 4139 TGTATTAATTTTGTATTAATTTT-----ATTTGTAATTTCTGCTTGTGAAAGAT 4190
|||||
DB 4232 ATGTGTAATTTTGTATTAATTTTCCATTTGAGCTGATGACTGCTTGTGTAAGAT 4291
|||||
OY 4191 TATGAAGAGTGAAGAAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGA 4249
|||||
DB 4292 TATGAAGAGTGAAGAAATTAATTAATTTGTAAGTGAAGTGAAGTGAAGTGAAGTGA 4348
|||||
OY 4250 ACTTTATATCAAAAAAATTTTAAAAA 4279
|||||
```


Db 4349 ACTTTCATGTGACTGGAAAAA 4378

RESULT 9
AAV32645
ID AAV32645 standard; cDNA: 4669 BP.
XX
AC AAV32645;
XX
DT 23-SEP-1998 (first entry)
XX
DE Human P glycoprotein (Pgp) cDNA.
XX
KW Human P glycoprotein; Pgp; multi-drug resistance; cancer;
KW UIC2 monoclonal antibody; mAb; cytotoxic; transmembrane efflux pump; ss.
XX
OS Homo sapiens.
XX
FH Key Location/Qualifiers
FT 5'UTR 1..424
FT CDS /tag-a
FT 425..4267
FT /tag-b
FT /Product- "Wild-type Pgp"
FT 4268..4669
FT /tag-c
XX
PN M09021325-A1.
XX
PD 22-MAY-1998.
XX
PF 17-NOV-1997; 97WO-US21214.
XX
PR 15-NOV-1996; 96US-0752447.
XX
PA (INGE-) INGENEX INC.
XX (UNIT) UNIT ILLINOIS FOUND.
XX
PI Mechtner E, Roninson IB;
XX
DR WPI: 1998-297930/26.
XX P-PSDB: AAM48997.
XX
PT Immunological reagent specific for P-glyco:protein - useful for
PT detecting multi-drug resistant cancer, isolating haematopoietic
PT cells and selective cell killing
XX
PS Claim 1; Fig 1A; 89pp; English.
XX
CC The present sequence represents the human P glycoprotein (Pgp)
CC cDNA which encodes the wild-type Pgp protein. Pgp is a transmembrane
CC efflux pump protein involved in multi-drug resistance of cancer cells.
CC The invention provides methods for developing and using immunological
CC reagents specific for certain mutant forms of Pgp and wild-type Pgp in
CC a conformation associated with substrate binding or in the presence
CC of ATP depleting agents. An example of the immunological reagent
CC is the UIC2 monoclonal antibody (mAb). mAb UIC2 specifically binds
CC to Pgp in a particular biochemical conformation and is capable of
CC inhibiting drug efflux from Pgp-expressing cells. The immunological
CC mammalian cells, including low level expression, particularly in cancer
CC cells to diagnose multi-drug resistance. The invention claims that
CC these immunological reagents are more specific than known reagents for
CC detecting Pgp and they also eliminate the need for costly and laborious
CC screening of Pgp inhibitors by cytotoxicity or dye exclusion methods.
XX
SQ Sequence 4669 BP; 1393 A; 894 C; 1130 G; 1252 T; 0 other;

Query Match 77.9%; Score 3333.2; DB 19; Length 4669;
Best Local Similarity 87.5%; Pred. No. 0;
Matches 3753; Conservative 0; Mismatches 493; Indels 44; Gaps 8;

OY 1 GGAGCGGAGGTGCGGATGATCTGAAGGAGCGCGTAAGGAGGTGCAGAGAAAGAACTT 60

Db 409 GGAGCGGAGGTGCGGATGATCTTGAAGGAGCGCGTAAGGAGGTGCAGAGAAAGAA 468

OY 61 CTGGAATAATGGGCAAAAAAGTAAAAAATGAGAGAAAGAAAGAAACCACTGTCAG 120

Db 469 CTTTAAAACTGAACATTAAGTGAAGAAAGTAAAGAGAAAGAAACCACTGTCAG 528

OY 121 CACGTTGCAATGTTTGGCTATTCAAATTTGGCTTGTATAGTGTATATGTTGGGGAC 180

Db 529 TGTATTTTCAATGTTTGGCTATTCAAATTTGGCTTGTATAGTGTATATGTTGGGGAC 588

OY 181 AATGCTCCCATCATCCATGAGCTGACCTCCCTCATGATGATGCTTGTGAAACAT 240

Db 589 TTGGGCTGCATCATGATGAGCTGACCTCCCTCATGATGATGCTTGTGAAACAT 648

OY 241 GACAGATAGCTTGGCAAAAGCAGAAATTTCAAGAAACAAACTTTCCAGTTATATTA 300

Db 649 GACAGATATCTTGGCAAAAGCAGAAATTTAGACATCTGATGT-----CAACATCAC 702

OY 301 TGAAGTATTTACGACATACATACATTTATCATCAACCATGAGGAGAAATGACAC 360

Db 703 TAATAGAGTATATCATATGATATACAGGTTCTTATGATCTGAGGAGAACGATACAC 762

OY 361 GTATGCTATTTATACAGTGGATGCTGCTGCTGCTGCTGCTGCTTACATCAGAT 420

Db 763 GTATGCTATTTATACAGTGGATGCTGCTGCTGCTGCTGCTGCTGCTTACATCAGAT 822

OY 421 TTCATTCGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 480

Db 823 TTCATTCGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 882

OY 481 TGTATTCATGCGACAGGAGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 540

Db 883 TGTATTCATGCGACAGGAGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 942

OY 541 CCGGCTCACAGACGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 600

Db 943 CGGACTTACAGATGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1002

OY 601 CTTTCACTCAATAGCAACATTTTACCGGTTTATATGATGCTGCTGCTGCTGCTGCTGCT 660

Db 1003 CTTTCACTCAATAGCAACATTTTACCGGTTTATATGATGCTGCTGCTGCTGCTGCTGCT 1062

OY 661 GCTAACCTTGTGATTTTGGGCAATCACCCTGCTTGTGCACTTGCACCGCATTCGTCG 720

Db 1063 GCTAACCTTGTGATTTTGGGCAATCACCCTGCTTGTGCACTTGCACCGCATTCGTCG 780

OY 721 AAGATACATCTTCACTTACTGATAAAGAACTCTTGGGCTATGCAAGCTGAGACAT 780

Db 1123 AAGATACATCTTCACTTACTGATAAAGAACTCTTGGGCTATGCAAGCTGAGACAT 1182

OY 781 AGCTGAGAGAGCTTGTAGCAGCAATCAGAACCTGATGCTTGTGAGGACAAAGAAAGA 840

Db 1183 AGCTGAGAGAGCTTGTAGCAGCAATCAGAACCTGATGCTTGTGAGGACAAAGAAAGA 1242

OY 841 ACTTGAAGGTACACAAAAATTTAGAGAGCTTAAGGAATTTGGATTTAAGAAAGCTAT 900

Db 1243 ACTTGAAGGTACACAAAAATTTAGAGAGCTTAAGGAATTTGGATTTAAGAAAGCTAT 960

OY 901 CAGGCGCAACATTTTCAATTTGTCGCTTGTATGATGATGATGATGATGATGATGATGAT 960

Db 1303 TACAGCAATATTTTCAATTTGTCGCTTGTATGATGATGATGATGATGATGATGATGAT 1362

OY 961 TTTCTGTATGAGGACCTCTTGTGCTCTCCGCTGAATATTTGATGATGATGATGATGAT 1020

Db 1363 CTTCGTGATGAGGACCTCTTGTGCTCTCCGCTGAATATTTGATGATGATGATGATGATGAT 1422

OY 1021 TGTCTTCTTTTGTATTAATTTGAGGCTTTTATGATGATGATGATGATGATGATGATGAT 1080

Db 1423 TGTATCTTTTGTATTAATTTGAGGCTTTTATGATGATGATGATGATGATGATGATGAT 1482

OY 1081 AGCATTTGCAAGCGAAGGAGGAGCTTATGAATCTTCAAGATTAATGACATAAACC 1140

Db 1483 AGCATTTCCAAATGCAGAGAGACGCTTATGAAATCTTCAGATTAATGTAATTAAGCC 1542
Oy 1141 AAGCATTTGACGCTATTTCCAGAGTGACATTAACCCAGATTAATTAAGGAAATTTGGA 1200
Db 1543 AAGTATTGACAGCTATTTCAGAGATGGGCAACAAACAGATTAATTAAGGAAATTTGGA 1602
Oy 1201 ATTCAAAAATGTTCACTTCAGTTACCTTCTCTGAAAAGAAAGTTAGATCTTAAGGGCT 1260
Db 1603 ATTCAAGAAATGTTCACTTCAGTTACCTTCGAAAAGAAAGTTAGATCTTGAAGGGCT 1662
Oy 1261 CAACCTGAAGCTTCAGAGTGGGACAGACATGGCGCTGTTGGGAACAGTGGCTCGGGA 1320
Db 1663 GAACCTGAAGCTTCAGAGTGGGACAGACGGGCGCTGTTGGGAACAGTGGCTCGGGA 1722
Oy 1321 GAGCAGCAGCGTGCAGCTGATGTCAGAGGCTCTATGACCCACAGATGGATGGTCTGAT 1380
Db 1723 GAGCAGCAACAGTCCAGCTGATGTCAGAGGCTCTATGACCCACAGAGGGATGGTCAAGT 1782
Oy 1381 TGATGACAGGACATTTAGACCAATTAATGTAAGCATCTTCGGGAAATTAAGTGGTGGT 1440
Db 1783 TGATGACAGGATATTAGACCAATTAATGTAAGCTTTCTAAGGAAATCATTTGTTGGT 1842
Oy 1441 GAGTCAGAGGCTGTTGTTGTTGGCCACACGATAGCTGAACATTCGCTATGGCCGGA 1500
Db 1843 GAGTCAGAGACCTGATTTGTTGGCCACACGATAGCTGAACATTCGCTATGGCCGGA 1902
Oy 1501 AAATGTCACCTGATGATGATGTAAGAAAGCTGTTAAGGAAGCAATGCCATGATTTAT 1560
Db 1903 AAATGTCACCTGATGATGATGTAAGAAAGCTGTCACAAAGAACCAATGCCATGATTTAT 1962
Oy 1561 CATGAACCTACCTAATAATTTGACACTGTGGTTGAGAGAGAGGGCCCAAGCTGATGG 1620
Db 1963 CATGAACCTGCTCAATAATTTGACACCTGTGGAGAGAGGGCCCAAGCTGATGG 2022
Oy 1621 TGGCAGAAACAGAGAAATCGCCATTTGCTCGGGCCCTGGTTGCAACCCCAAGATTTCT 1680
Db 2023 TGGGAGAGAGAGAGATCGCCATTTGCTCGGGCCCTGGTTGCAACCCCAAGATTTCT 2082
Oy 1681 GCTGGATGAGCAACGTCACGCTGAGCACTGAAAGTGAAGCAGTGGTTCAAGTGGCT 1740
Db 2083 GCTGGATGAGGCAAGTCAAGCTTGGACACAAAGGCAAGTGGTTCAAGTGGCT 2142
Oy 1741 GGAATAGCCAGAAAGCCGAGCTACACATTTGATGATGATGATGATGATGATGATG 1800
Db 2143 GGAATAGCCAGAAAGGTCGAGCACTGATGATGATGATGATGATGATGATGATG 2202
Oy 1801 TAATGCGATGTCATGCTGCTGTTTGAATGATGATGATGATGATGATGATGATG 1860
Db 2203 TAATGCGATGTCATGCTGCTGTTTGAATGATGATGATGATGATGATGATGATG 2262
Oy 1861 TGAATCATGAAAGAGAGGCAATTTACTTCAAACTGTGCAAAATGCAAGACAGAGAA 1920
Db 2263 TGAATCATGAAAGAGAGGCAATTTACTTCAAACTGTGCAAAATGCAAGACAGAGAA 2322
Oy 1921 TGAATGATGATGAAATATCCACTGTGATTCAAATGAAATGAAATGATGCTTGAAT 1980
Db 2323 TGAATGATGATGAAATATCCACTGTGATTCAAATGAAATGAAATGATGCTTGAAT 2382
Oy 1981 GTCTCAGAAAGATTCAGGCTCCAGTATTAATAAAGAAATCAACGCGAGATATCA 2040
Db 2383 GTCTCAGAAATGATTCAGATTCAGATTCATTAATAAAGAAATCAACGCGAGATATCA 2442
Oy 2041 TGCACCAAGGCCAAGACAGAAAGCTTGTACAAAGAGAGACTTGATAGATGATACC 2100
Db 2443 TGCACCAAGGCCAAGACAGAAAGCTTGTAGTACCAAGAGGCTGTGAGAAAGTATACC 2502
Oy 2101 TCCAGTTTCCTTGGAGAGATTCGAAGCTGAACCTCACTGAATGGCTTATTTGGT 2160
Db 2503 TCCAGTTTCCTTGGAGAGATTCGAAGCTGAATTTAACTGAATGGCTTATTTGGT 2562
Oy 2161 TGGTATATTTTGGCTATTAATAAGAGAGGCTGCAACAGACTTTTCAATAATATTTTC 2220
Db 2563 TGGTATATTTTGGCTATTAATAATGAGGCTGCAACAGACTTTTCAATAATATTTTC 2622

Oy 2221 AAGATTTATAGGATCTTTACCCAGATGAGATCTGAAACAAACAGACAGATATGTA 2280
Db 2623 AAGATTTATAGGATTTTATACAGATTTGATGATCTGAAACAAACAGACAGATATGTA 2682
Oy 2281 CATGTTTCTGATGATGTTTCTAGCTTGGAAATTAATTTCTTTATTAATTTTCTTCA 2340
Db 2683 CTGTTTCTGATGATGTTTCTAGCTTGGAAATTAATTTCTTTATTAATTTTCTTCA 2742
Oy 2341 GGGCTTCACATTTGGCAAGCTGGGAGATCTCTCAAGGGGGCTGATCATGTTT 2400
Db 2743 GGGCTTCACATTTGGCAAGCTGGGAGATCTCTCAAGGGGGCTGATCATGTTT 2802
Oy 2401 CAATTCAGCTGAGACAGATATGATGATGATGATGATGATGATGATGATGATG 2460
Db 2803 CAATTCAGCTGAGACAGATATGATGATGATGATGATGATGATGATGATGATG 2862
Oy 2461 ATTGACAAACAGCTTCCCAATGATGCGGCTCAAGTTAAAGGGGCTATAGTTCCAGCT 2520
Db 2863 ATTGACAAACAGCTTCCCAATGATGCGGCTCAAGTTAAAGGGGCTATAGTTCCAGCT 2922
Oy 2521 TGGTGTATTAACCAAGATATGATGATGATGATGATGATGATGATGATGATG 2580
Db 2923 TGGTGTATTAACCAAGATATGATGATGATGATGATGATGATGATGATGATG 2982
Oy 2581 TGGTGGCAATTAACACTTTTACTTTAGCATTTGATGATGATGATGATGATGATG 2640
Db 2983 TGGTGGCAATTAACACTTTTACTTTAGCATTTGATGATGATGATGATGATGATG 3042
Oy 2641 TGGTGAATTAATTAATGTTGCTGACAGACACTGAAAGATTAAGAAAGCTAGAGAG 2700
Db 3043 TGGTGAATTAATTAATGTTGCTGACAGACACTGAAAGATTAAGAAAGCTAGAGAG 3102
Oy 2701 TGGGAAGATTTCTACAGAAAGCATGAAAACTTCCGAACTGTTGTTTCTTACTGCGGA 2760
Db 3103 TGGGAAGATTTCTACAGAAAGCATGAAAACTTCCGAACTGTTGTTTCTTACTGCGGA 3162
Oy 2761 GCAGAGTTTGAATTAATGATGATGATGATGATGATGATGATGATGATGATG 2820
Db 3163 GCAGAGTTTGAATTAATGATGATGATGATGATGATGATGATGATGATGATG 3222
Oy 2821 GAAAGCAGACATCTTCCGGGCTCATTTTCTATCCACCCAGGCAATGATTTTCC 2880
Db 3223 GAAAGCAGACATCTTGGAAATTAATTTCTTCCACCCAGGCAATGATTTTCC 3282
Oy 2881 TCTGCTGCTTCCGCTTGGTGGCTTGGTGGCTTGGTGGCTTGGTGGCTTGG 2940
Db 3283 TCTGCTGCTTCCGCTTGGTGGCTTGGTGGCTTGGTGGCTTGGTGGCTTGG 3342
Oy 2941 TGTCTTTTGGTATTTCTCAGTATTTGCTTTGTTGTTGTTGTTGTTGTTGTTG 3000
Db 3343 TGTCTTTTGGTATTTCTCAGTATTTGCTTTGTTGTTGTTGTTGTTGTTGTTG 3402
Oy 3001 ATTGCTCTGATGATGCAAAAGCCAAATATACACACCCAGGCAATCATGATCAT 3060
Db 3403 ATTGCTCTGATGATGCAAAAGCCAAATATACACACCCAGGCAATCATGATCAT 3462
Oy 3061 AAAAAGCCCTGATTTGACAGCTACAGCCCTCAGGCCCTCAAGCCAAATACGTTGAAG 3120
Db 3463 AAAAAGCCCTTATTTGACAGCTACAGCCAGGAGCCCTTAAGGCCGAACATTTGAAG 3522
Oy 3121 AAATGTCATTTAATGAGTCTGTTCAACTATCCCACTGCAACAGACATCCCGCT 3180
Db 3523 AAATGTCATTTGATGAGTCTGTTCAACTATCCCACTGCAACAGACATCCCGCT 3582
Oy 3181 CCAGGGCTGATGCTGAGGTTGAAGAGGGCCAGAGCTGGCCCTGATGATGATG 3240
Db 3583 TCAGGGCTGATGCTGAGGTTGAAGAGGGCCAGAGCTGGCCCTGATGATGATG 3642
Oy 3241 CTGTTGGAGAGAGCAGATTTGAGCTCTAGAGGCTTCTATGACCCCTTGGCTGGTTC 3300
Db 3643 CTGTTGGAGAGAGCAGATTTGAGCTCTAGAGGCTTCTATGACCCCTTGGCTGGTTC 3702

OY	1201	ATTCCAAATATGTTCACTTCAGTTACCTTCTCTCGAAAANAATTAAGTCTTAAAGGTC	1280
Db	1603	ATTCCAAATATGTTCACTTCAGTTACCTTCTCGAAAANAATTAAGTCTTAAAGGTC	1662
OY	1261	CAACCTGAAGGTTCTGAGTGGGCGACACATGGCGCTGGTTGGAAACGTTGGCGGAA	1320
Db	1663	GAACCTGAAGGTTGGAGATGGGCGACACGCTGGCGCTGGTTGGAAACGTTGGCGGAA	1722
OY	1321	GACACACGACCGTGGAGTATCGAGAGGCTCATACCCACAGATGGATGGTCTGTAT	1380
Db	1723	GACACACAGTCCGAGTGGATCGAGGCTCATATACCCACAGAGGGATGGTCAAGT	1782
OY	1381	TGATGGACAGGACATATGAGACCTTAATATGAAAGCATTTGGGAATTTCTGTTGGT	1440
Db	1783	TGATGGACAGGATATTAGACCATTAATATGAAGTTTCTACGGAAATCATTTGGTGGT	1842
OY	1441	GAGTCAGAGGCGCTGTGTGTGTGGCACACGATACGTGAATAACATTGGCTATGGCCGGA	1500
Db	1843	GAGTCAGGAACCTGTATGTGTGTGGCACACGATACGTGAATAACATTGGCTATGGCCGGA	1902
OY	1501	AAATGTCACACATGATGATGATGGAAGACCTTTAAGAGGCCATGCTATGATTTAT	1560
Db	1903	AAATGTCACACATGATGATGGAAGACCTGTCAAGGAAACCATGCTATGATGATTTAT	1962
OY	1561	CATGAACACTACTATTAATTTGACACTGTGGTTGGAAGAGAGAGGGCCAGCTGAGTGG	1620
Db	1963	CATGAACACTCCTCTAAATTTGACACCTGGTTGGAGAGAGAGGGCCAGTGTAGTGG	2022
OY	1621	TGGACAGAAACAGAAATACGCAATGCTGTGGGCGCTGGTTGCAACCCCAAGATCTTCT	1680
Db	2023	TGGGCGAAGACGAGATACGCAATGCAAGTCCCTGGTTGCAACCCCAAGATCTCTCT	2082
OY	1681	CGTGGATGAGGCACTGCTGAGTCTGAGACCTGAAGATGAACAGTGGTTAGGTGGCCCT	1740
Db	2083	CGTGGATGAGGCACTGCTGAGTCTGAGACAGAAAGCAAGAGAGGTATAGGTGGCTCT	2142
OY	1741	GGATTAAGGCGAAGAAAGGCGGACTACCATTTGATAGCTCATGCTTTGTCTACAGTTGG	1800
Db	2143	GGATTAAGGCGAAGAAAGGCTCGACACCATTTGATGATCTCATGCTTTGTCTACAGTTGG	2202
OY	1801	TAATGCCGATGTCATTTGCTGGTTTGTGATGATGAGACTCATTTGGAGAAAGCAATCATGA	1860
Db	2203	TAATGGCTGACGTCAATCGCTGGTTTGTGATGATGAGACTCATTTGGAGAAAGCAATCATGA	2262
OY	1861	TGAAGTCATGAAGAAAGGAAGGCGCATTTACTTCAACCTGTCCACATTCGACAGCAAGAGGAA	1920
Db	2263	TGAAGTCATGAAGAAAGGAAGGCGCATTTACTTCAACCTGTCCACATTCGACAGCAAGGAA	2322
OY	1921	TGAAATTTGATTAAGAAATGCGCACTGGTGAATTCGAAAAGTGAAGTACGCTTGGAAAT	1980
Db	2323	TGAAGTTGATTAAGAAATGAGCTGTGATGATTCGAAAAGTGAAGTATGATGCTTGGAAAT	2382
OY	1981	GTCGTCCAAAGATTTGAGGCTCCAGTTTATTAATAAAGAAAGATCAACTCGCAGGATATACA	2040
Db	2383	GTCCTTCAATGATTTCAATCATCCAGTCTAATTAAGAAAAAGATCAACTCGTGGAGTGTCCG	2442
OY	2041	TGCACACCAAGGCCAAGACAGAAAGCTTGGTGTCAAAAAGAGACTGTGATAGCATGTAC	2100
Db	2443	TGCATACCAAGGCCAAGACAGAAAGCTTGTGTGTCCAAAAGGCTCTGGATGAAGTATAC	2502
OY	2101	TTCAGATTTCTCTGAGAGATTCGGAAGCTGGAAGCTCACTGAATAGGCGCTAATTTGTGTGT	2160
Db	2503	TTCAGATTTCTCTTTGAGAGATATTGAAGCTAAATTTAATCTGAATAGGCGCTAATTTGTGTGT	2562
OY	2161	TGGTATATTTTGTGCTATATTAAACGAGAGCTCGACACACGATTTTCAATTAATATTTC	2220
Db	2563	TGGTATATTTTGTGCTATTAATTAATGAGAGGCTCGACACACGATTTTCAATTAATATTTC	2622
OY	2221	AAGGATTTATAGGATCTTTACCCGACATAGAGATCCTAACAACAAAGCAGAGATAGTAA	2280
Db	2623	AAGGATTTATAGGATTTTACAGAAATATTATGATCCTGAACAACAAAGCAGAGATAGTAA	2682
OY	2281	CAGTTTCTGTGATGTTCTTGAGGCGCTGGGAATTTTCTTTTATTCATTTTCTCTCA	2340

Db	3763	GGGCGATGCTGTCCAGGAGGCCACCTCCTGTTTGACTGCAGAGATTGCTGGAACAACATTGGCTTA	3822
OY	3421	TGGAGACAAACAGCCGGGTGCTATACATAGAAAGATTTATGCAAGCAGCCAGAGAGGCCAA	3480
Db	3823	TGGAGACAACAGCCGGGTGCTGTACACAGAAAGATGTGAGAGGCCAGCAAGAGAGGCCAA	3882
OY	3481	CATACCACTTTCATCGACGACACTCCCTGGAATTAACAACACAGAGTGTAGAGCAAAAG	3540
Db	3883	CATACATGCTTTCATCGAGTCACTGCTCTAATTAATATACACTAAAGTAGAGAGCAAAAG	3942
OY	3541	AACCAAGCTCTGTGTGGCCAGCAAAAGAGGCAATGCGATAGTCCGCTCTTGTTAGACA	3600
Db	3943	AACATGACTCTGTGTGGCCAGCAAAAGAGGCAATGCGATAGTCCGCTCTTGTTAGACA	4002
OY	3601	GCCATCATTTTTCCTTTTGGATGAAGCAATCACTGCTGTGATCAGAAAGTGAAGAGT	3660
Db	4003	GCCATCATTTTTCCTTTTGGATGAAGCAACGTCAGCTCTGTGATCAGAAAGTGAAGAGT	4062
OY	3661	TGTCCAAAGAGCCCTGTGACAAAGCCAGAGAGGCCGACCTGTGATGTGATCGCCACCG	3720
Db	4063	TGTCCAAAGAGCCCTGTGACAAAGCCAGAGAGGCCGACCTGTGATGTGATCGCCACCG	4122
OY	3721	CTTGTCCACCATTCACGAATGCAAGATTTAAATAGTGTGTTCAGAAATGGCAAGTCAAGA	3780
Db	4123	CGTGTCCACCATTCACGAATGCAAGATTTAAATAGTGTGTTCAGAAATGGCAAGTCAAGA	4182
OY	3781	GCATGGCACACATCAACAGCGTGTGTGGCCGCAAGAGGCATCTATTTTCCATGTGCAGTG	3840
Db	4183	GCATGGCACACATCAACAGCGTGTGTGGCCGCAAGAGGCATCTATTTTCCATGTGCAGTG	4242
OY	3841	CCAGGCTGTGAGCAAAAGCGGTAGTGAACGTGTGGCCATATGAGCTGTAAATATTTTAAAT	3900
Db	4243	CCAGGCTGTGAGCAAAAGCGGTAGTGAACGTGTGGCCATATGAGCTGTAAATATTTTAAAT	4302
OY	3901	ATTGTGTGTTAAACATGSCATTTAATCAAGTTAAAGGTGAGCACTTACTGGAAGAACT	3960
Db	4303	ATT--TGTTTAGATATATACATTATTAACAAGTTAA--AGCAAAACACTTACAGAAAT	4355
OY	3961	ATGTAGAACATACCTGTTTAAACATTTCTGTGTCACTGAAGATCAATCCACCAAGTTGAG	4020
Db	4356	ATGTAGAACATACCTGTTTAAACATTTCTGTGTCACTGAAGATCAATCCACCAAGTTGAG	4403
OY	4021	AGCTTCAGATTTTATATATTAAGGAACAA--AGAAACATTTATCTGATGAATTAAT	4078
Db	4404	----AGAGACTTCGTAATTTAAAGCAAGATGTGAGACATCAATCAAGTGAAGAAAT	4458
OY	4079	ACTGTGTATATTCATTAATTAATTAATATAGATAATTCAAAGTAGATTTTGTAAATAAT	4138
Db	4459	CATAGTTTAAACGCAATTAATTAATTTATTAACAAGATTAAGTAGATTTTAAAGATTA	4518
OY	4139	TGTATATTTTGTGTATATTTT-----ATTGTACTTACTGCTTGTCTGAAGAT	4190
Db	4519	ATGTGTAATTTTGTGTATATTTTCCCAATTTGGACTGTGAACGATGCTGCTTGAAGAT	4578
OY	4191	TATAGAACTGTGAATAAAGTACTG--AATGTTTGAATTAAGTGTCTGCTATTAATTAAC	4249
Db	4579	TATAGAACTGTGAATAAAGTATGAATTTGCAATTAAGTCT--TCTATATATTAAC	4635
OY	4250	ACTTTTATATCAAAAAAAAAAAAAAAAAAAAA 4279	
Db	4636	ACTTTCATGTGAAAAAAAAAAAAAAAAAAAAA 4665	
RESULT 11			
ID	AA294738	standard; cDNA; 4646 BP.	
XX	AA294738;		
XX	01-AUG-2000	(first entry)	
XX	Human ATP binding cassette ABCB1 (MDR1) cDNA,		
XX			

1441 GAGTCAGGAGCCTGTTGTTGCCACACGATAGCTGAAAACAATTGGCTATGGCCGCA

Db 1843 GAGTCAGGAACCTGATATTGTTGGCCACCACGATAGCTGAAGAAACATTCGCTATGGCCCTGA 1902
QY 1501 AAATGTCACATGATGATGATGAGAAAGCTGTTAAGAAAGCCATGCCATGATTTTAT 1560
Db 1903 AAATGTCACATGATGATGATGAGAAAGCTGTCAGAAAGCCAAATGCCATGATTTTAT 1562
QY 1561 CATGAACCTACCTAATTAATTTGACACTTCGTTGGAGAGAGAGGGGCCAGCTGAGTGG 1620
Db 1963 CATGAACCTGCTCAATTAATTTGACACCCCTGGTTGGAGAGAGAGGGGCCAGCTGAGTGG 2022
QY 1621 TGGACAGAAAGAGATCCGCAATGCTCGGCCCTTCGTTCCGACCCCAAGATTTCTCT 1680
Db 2023 TGGGCAAGAGAGAGATCCGCAATGCACTGCCCTGTTCCGACCCCAAGATTCCTCT 2082
QY 1681 GCTGGATGAGGCAACCTCAGCTCTGACACTGAAAGTGAAGAGAGAGTTCAGGTGGCCCT 1740
Db 2083 GCTGGATGAGGCAACCTCAGCTCTGACACACAGAAAGCGAAGAGAGTTCAGGTGGCCCT 2142
QY 1741 GGATGAAGGCGCAAAAAGGCCGGACCTACATTCGATAGCTCATCGTTGCTACAGTTTCG 1800
Db 2143 GGATGAAGGCGCAAAAAGGTCGGACACCAATGATAGCTCATCGTTGCTACAGTTTCG 2202
QY 1801 TAAATCCGATGCTATTCGTTGTTTGTATGATGAGTCAATTCGTTGGAGAAAGAAATCATGA 1860
Db 2203 TAAATGCTGACGTCATCGTGGTTTGATGATGAGTCAATTCGTTGGAGAAAGAAATCATGA 2262
QY 1861 TGAACCTCATGAAGAGAAAGGCAATTCCTCAAACTTCGCAATGCAAGAGAAAGAA 1920
Db 2263 TGAACCTCATGAAGAGAAAGGCAATTCCTCAAACTTCGCAATGCAAGAGAGAA 2322
QY 1921 TGAATTTGATTAAGAAATGCACTGCTGATGATCCAAAGTGAAGTGAATGCTTGAAT 1980
Db 2323 TGAATTTGATTAAGAAATGCACTGATGATCCAAAGTGAAGTGAATGCTTGAAT 2382
QY 1981 GTCCTCAAAAAGATTAAGGCTCCAGTTTAATTAATAAAGATCAACTGCAAGAGATATGA 2040
Db 2383 GTCCTCAAAAAGATTAAGGCTCCAGTTTAATTAATAAAGATCAACTGCAAGAGATATGA 2442
QY 2041 TGCACACAAAGGCCAAGACAGAAAGCTGGTAAAGAAAGAGACTTCGATGAATGATAC 2100
Db 2443 TGCATCACAAAGGCCAAGACAGAAAGCTAGTACAAAGAGAGCTCTGATGAATGATAC 2502
QY 2101 TCCAGTTCCCTTCGAGAGATTCGAAGCTGAACCTGAATGGCTTATTTTGGT 2160
Db 2503 TCCAGTTCCCTTCGAGAGATTCGAAGCTGAATGATGAAGCTAATTTAATGCTTATTTTGT 2562
QY 2161 TGTATATTTTGTCTATTTAATTAACGAGGCCCTGCAACGCAATTTTCAATATATTTTC 2220
Db 2563 TGTGTATTTTGTCTATTTAATTAATGAGGCTGCAACGCAATTTTCAATATATTTTC 2280
QY 2221 AAGATTAATAGGATCTTTAACCAGATGAGATCCCTGAACAAAGCAGAGATATGTA 2280
Db 2623 AAGATTAATAGGATCTTTAACCAGATGAGATCCCTGAACAAAGCAGAGATATGTA 2682
QY 2281 CATGTTTCTGATATTTGTTTATGCTTGAATATTTCTTTTATTTATTTCTCTCA 2340
Db 2683 CATGTTTCTGATATTTGTTTATGCTTGAATATTTCTTTTATTTATTTCTCTCA 2742
QY 2341 GGGCTTCACATTTTGGCAAAAGCTGGGAGATCCCTAAGCGGCTTCATATCATGTTT 2400
Db 2743 GGGCTTCACATTTTGGCAAAAGCTGGGAGATCCCTAAGCGGCTTCATATCATGTTT 2802
QY 2401 CAGATTCATGCTGAGCAGAGATGTCAGCTGGTTCATGACCCCTAAAAACACACTGAGC 2460
Db 2803 CAGATTCATGCTGAGCAGAGATGTCAGCTGGTTCATGACCCCTAAAAACACACTGAGC 2862
QY 2461 ATTGACAAACAGGCTTGCCAAATGATCGGCTCAAGTTAAAGGGCTATAGTTCCAGCT 2520
Db 2863 ATTGACAAACAGGCTTGCCAAATGATCGGCTCAAGTTAAAGGGCTATAGTTCCAGCT 2922
QY 2521 TCGTGCATTTACCGAGATATGCAATCTTGGACAGGCAATTTATTCCTTATCTTAATCTA 2580
Db 2923 TCGTGCATTTACCGAGATATGCAATCTTGGACAGGCAATTTATTCCTTATCTTAATCTA 2982

QY 2581 TCGTGGCAATTAACACTTTTACTTCTTAGCAATTTGTAACCATCATTTGCAATGAGGAG 2640
Db 2983 TCGTGGCAATTAACACTTTTACTTCTTAGCAATTTGTAACCATCATTTGCAATGAGGAG 3042
QY 2641 TGTGAAATGAAATGTTTGTCTGAGCAAGCAGCTGAAAGATTAAGAAAGCTAGAGAGC 2700
Db 3043 TGTGAAATGAAATGTTTGTCTGAGCAAGCAGCTGAAAGATTAAGAAAGCTAGAGAGC 3102
QY 2701 TGGCAAGATTTCTACAGAAAGCCATCGAAACCTTCGAACTGTGTTCTTACATCGGA 2760
Db 3103 TGGCAAGATTTCTACAGAAAGCCATCGAAACCTTCGAACTGTGTTCTTACATCGGA 3162
QY 2761 GCAGAAATTTGAATATGATATGATACAGAGTTTGGCAATGATACATACAGAACTTTGAG 2820
Db 3163 GCAGAAATTTGAATATGATATGATACAGAGTTTGGCAATGATATGATATGATATGAT 3222
QY 2821 GAAAGCAGACATCTTCGGGGTCTCATTTCTATACCCGAGCAATGATATTTTCTTA 2880
Db 3223 GAAAGCAGACATCTTCGAAATTAATTTCTTACACCCAGGCAATGATATTTTCTTA 3282
QY 2881 TCGTGCATTTTCCGGTTTGGTGGCTACTTGTGTCGCAATGATGATGATGATGATG 2940
Db 3283 TCGTGCATTTTCCGGTTTGGTGGCTACTTGTGTCGCAATGATGATGATGATGATG 3342
QY 2941 TGTCTTTTGGATTCACGCTATTTGCTTTGGTGCATGGCAGTGGGGGAGCTGACGTC 3000
Db 3343 TGTCTTTTGGATTCACGCTATTTGCTTTGGTGCATGGCAGTGGGGGAGCTGACGTC 3402
QY 3001 ATTGTCCTGCTATTCGCAAAAGCCAAAGTATCAGACGCCAGCTCATCATGATCATGA 3060
Db 3403 ATTGTCCTGCTATTCGCAAAAGCCAAATATCAGACGCCAGCTCATCATGATCATGA 3462
QY 3061 AAAAAGCCCTGATTTGACAGCTACAGCCCTCAGCGGCTCAGGCAATATGCTGGAAG 3120
Db 3463 AAAAAGCCCTGATTTGACAGCTACAGCCCTCAGCGGCTCAGGCAATATGCTGGAAG 3522
QY 3121 AAATGTCATTTAATGAGTTCGCTGCTTCAACTATCCCATGCAACAGATCCCGTGT 3180
Db 3523 AAATGTCATTTAATGAGTTCGCTGCTTCAACTATCCCATGCAACAGATCCCGTGT 3582
QY 3181 CCAGGGGCTGAGCCCTCAGAGTGAAGAGGCGCAGAGCCCTGCTGAGTATGCACTGG 3240
Db 3583 TCAGGAGCTGAGCCCTGAGAGTGAAGAGGCGCAGAGCCCTGCTGAGTATGCACTGG 3642
QY 3241 CTGAGGAAAGGCAAGTTCGCTTGAAGCTCTGAGAGCCCTTATGACCCCTGCTGCTTC 3300
Db 3643 CTGAGGAAAGGCAAGTTCGCTTGAAGCTCTGAGAGCCCTTATGACCCCTTGGCAGGAA 3702
QY 3301 AGTCTAATTTGATGCGAAAGATTAAGGACCTGATATGCTGCTGAGCAGACACT 3360
Db 3703 AGTCTCTTTGATGCGAAAGATTAAGGACCTGATATGCTGCTGAGCAGACACT 3762
QY 3361 GGGCATTCGCTTCAGAGAGCCCATCTGTTTGAAGTACAGACATTTGCTGAAACATTTGCTTA 3822
Db 3763 GGGCATTCGCTTCAGAGAGCCCATCTGTTTGAAGTACAGACATTTGCTGAAACATTTGCTTA 3882
QY 3421 TGGAGCAACAGCGGGGCTGTATACATGAGAGATTAATGCAAGGAGCCAAAGAGGCCAA 3480
Db 3823 TGGAGCAACAGCGGGGCTGTATACATGAGAGATTAATGCAAGGAGGCCAAAGAGGCCAA 3882
QY 3481 CATTACCACTTCATGAGACACTCCCTGAGAAATACAAACAGAGATGAGAGCAAAAG 3540
Db 3883 CATTACCACTTCATGAGACACTCCCTGAGAAATACAAACAGAGATGAGAGCAAAAG 3942
QY 3541 AACCCAGCTCTCTGGGCGCAAGAAACAGGCAATTCATGCTGCTGCTTGTATGACA 3600
Db 3943 AACCCAGCTCTCTGGGCGCAAGAAACAGGCAATTCATGCTGCTGCTTGTATGACA 4002
QY 3601 GCTCATATTTTCTTTTGGATGAGAGTACATCAGCTCTGGATACAGAAAGTAAAAAGT 3660
Db 4003 GCTCATATTTTCTTTTGGATGAGAGTACATCAGCTCTGGATACAGAAAGTAAAAAGT 4062

```

OY 3661 TGTCCAGAGGCGCTGGACAAAGCCAGAGAGCGCCGACCTGCATTGTGATGCCACCG 3720
    |||||
DB 4063 TGTCCAGAGGCGCTGGACAAAGCCAGAGAGCGCCGACCTGCATTGTGATGCCACCG 4122
    |||||
OY 3721 CTTGTCCAGCATCCAGAAATGCAGATTAAATAGTGTTTCAGAAATGGCAAGCAAGCA 3780
    |||||
DB 4123 CCGTCTCAGCATCCAGAAATGCAGACTTAATAGTGTTTCAGAAATGGCAAGCAAGCA 4182
    |||||
OY 3781 GGTATGGCAGCATCCAGAGAGCGCTGGCCAGAAAGCATATTTTCCATGTCAGTGT 3840
    |||||
DB 4183 GCATGGCAGCATCCAGAGAGCGCTGGCCAGAAAGCATATTTTCCATGTCAGTGT 4242
    |||||
OY 3841 CCAGGCTGGAGCAAGCGCTAGTGAAGTGGCCATAGAGCTGTAAATATTTTAAAT 3900
    |||||
DB 4243 CCAGGCTGGAGCAAGCGCCAGTCTGACTGTATGAGTGTAAATATCTTTTAAAT 4302
    |||||
OY 3901 ATTTGCTTAAACATGCGCATTTAATCAAGTTAAAGGTAGACACTTACGAAAAACT 3960
    |||||
DB 4303 ATT--TGTATGATATGACATTTTATTCAAAGTTAA--AGCAAGCACTTACGAAAT 4355
    |||||
OY 3961 ATGAGAACTACCTGTTTAAACATTTCTGCTGCACTGAAGTCAATCCACCAAGTTCAG 4020
    |||||
DB 4356 ATGAGAGGATATCGTTTAAACATTTCTGCTGCACTGAAGTCAATCCACCAAGTTCAG 4403
    |||||
OY 4021 AGCTTCAAGTTTAAATTAAGAAACCA--AAGAACTATTCATGATGAATTAAT 4078
    |||||
DB 4404 ----AGAGACTTGTATTAAGAAACCAAGCAAGTCAATCCACCAAGTTCAG 4458
    |||||
OY 4079 ACTGCTGTAATTCATTAATTAAGAAATTAAGATTAATTCATTAATTTGTTTAAAT 4138
    |||||
DB 4459 CATGATTTAAACATCATTAATTAATTAAGAAATTAAGATTAATTTGTTTAAAT 4518
    |||||
OY 4139 TGTATATTTTGTGTTTAAATTTT-----ATTGTAATCTACTGCTTGAAGAT 4190
    |||||
DB 4519 ATGTGTAATTTGTTTAAATTTTCCCATTTGCACTGCACTGCACTGCTTGAAGAT 4578
    |||||
OY 4191 TATGAGAACTGTTAAAGTACTG--AATGTTGAATTAAGTCTAGCTAATTAATTAAT 4249
    |||||
DB 4579 TATGAGAACTGTTAAAGTACTG--AATGTTGAATTAAGTCTAGCTAATTAATTAAT 4635
    |||||
OY 4250 ACTTTTATAT 4259
    |||||
DB 4636 ACTTTCATGT 4645
    |||||

```

RESULT 12

AAD38994 standard: DNA; 4646 BP.

AAD38994;

23-SEP-2002 (first entry)

Human mdr1 gene.

Human; haematologic malignancy; multidrug resistance; MDR; SUMO-1; hypoxia inducible factor-1; small ubiquitin-like-modifier; HIF-1; lymphoid disorder; chronic lymphoproliferative disorder; lymphoma; myeloid disorder; lymphocytic leukaemia; thrombocythemia; myeloma; angioimmunoid myeloid metaplasia; myeloid leukaemia; gene therapy; polycythemia vera; hypoxia responsive element; HRE; gene; ds.

Homo sapiens.

Location/Qualifiers

Key 425..4267

/*tag= a /product= "Human MDR protein"

W0200234291-A2.

02-MAY-2002.

25-OCT-2001; 2001WO-US49856.
26-OCT-2000; 2000US-243542P.
(BIGHAM) BRIGHAM & WOMENS HOSPITAL INC.
Colgan SP;
WPI, 2002-471427/50.
P-PSDB: AAE24211.
Treating a subject (at risk of) having a hematologic malignancy or multidrug resistance, e.g. lymphoma or myeloma, by administering hypoxia inducible factor 1 binding molecules or small ubiquitin-like-modifier-1 binding molecules -
Disclosure; Page 53-55; 92pp; English.

The invention relates to a method of treating a subject having or at risk of developing a haematologic malignancy or multidrug resistance (MDR). The method involves administering hypoxia-inducible factor-1 (HIF-1) binding molecules or small ubiquitin-like-modifier (SUMO)-1 binding molecules or HIF-1-SUMO-1 complex modulators. mdr1-hypoxia responsive element (HRE) binding molecules or antisense nucleic acid molecules and SUMO-1 binding molecules or antisense nucleic acid molecules are useful for treating a subject having or at risk of developing hematologic malignancy or MDR (e.g. a lymphoid disorder or a myeloid disorder). The lymphoid disorders include lymphocytic leukaemia or chronic lymphoproliferative disorders e.g. lymphoma, myeloma or chronic lymphoid leukaemia. The myeloid disorders include chronic or acute myeloid leukaemia or polycythemia vera. The invention is used in gene therapy. The present sequence is human mdr1 gene.

Sequence 4646 BP; 1371 A; 892 C; 1129 G; 1254 T; 0 other:

Query Match 77.7%; Score 3324.4; DB 24; Length 4646;
Best Local Similarity 87.6%; Pred. No. 0;
Matches 3740; Conservative 0; Mismatches 486; Indels 44; Gaps 8;

```

OY 1 GGAGCGCGAGGTGCGGATGATCTGAAGAGCGCCGTAAGCGGAGGCAAGAAACTT 60
DB 409 GGAGCGCGAGGTGCGGATGATCTGAAGAGCGCGGATGAGGAGCAAGAAAGAA 468
    |||||
OY 61 CTGGAATATGGCAAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 120
    |||||
DB 469 CTTTAACTGACATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 528
    |||||
OY 121 CAGCTTGCATGTTGCTATTCATTAATTTGCTGATAGTGTATGTTGGGGGAC 180
    |||||
DB 529 TGTATTTCAATGTTGCTATTCATTAATTTGCTGATAGTGTATGTTGGGGGAC 588
    |||||
OY 181 AATGCTGCATCATCATGAGATGAGCTCCCTCATGATGCTGTTTGGAAACAT 240
    |||||
DB 589 TTTGCTGCATCATCATGAGATGAGCTCCCTCATGATGCTGTTTGGAAACAT 648
    |||||
OY 241 GACAGATACCTTTCGAATGAGGATTTCAAGAAACAACTTTCCAGTTATTAATTA 300
    |||||
DB 649 GACAGATACCTTTCGAATGAGGATTTCAAGAAACAACTTTCCAGTTATTAATTA 702
    |||||
OY 301 TGAAGATTTACAGCAATACACAACTTTCATCAACATCTGAGAGGAATGACCAC 360
    |||||
DB 703 TAAATGAGATGATATCAATGATACAGGTTCTTCATGATCTGAGAGGAATGACCAC 762
    |||||
OY 361 GTATGCTATTTATTAAGTGGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 420
    |||||
DB 763 GTATGCTATTTATTAAGTGGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 822
    |||||
OY 421 TTCAATTCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 480
    |||||
DB 823 TTCAATTCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 882
    |||||
OY 481 TGTATATCATGAGCAGAGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 540
    |||||

```

Db 883 TCGTAAATGCGACAGAGATAGGCTGTTGATGTGACAGATGTTGGAGGCTTAAACAC 942
QY 541 CCGGCTCAGACAGCTGCTCCAAATCAATGAAGAAATGGCCACAAAATTTGAAATGTT 600
Db 943 CCGACTTCAGATGATGCTCTAAGATTAAGAAATGATGATGACAAAATTTGAAATGTT 1002
QY 601 CTTTCACTCAATAGCAACATTTTACCGGGTTTATAGTGGGGTTTACAGCTGTTGAA 660
Db 1003 CTTTCACTCAATAGCAACATTTTACCGGGTTTATAGTGGGGTTTACAGCTGTTGAA 1062
QY 661 GCTAACCCCTTGATTTTGGCCATCAGCCCTGTTTGGACCTTTCAGCCGCCATCGGGC 720
Db 1063 GCTAACCCCTTGATTTTGGCCATCAGCCCTGTTTGGACCTTTCAGCTGCTGCTGGCC 1122
QY 721 AAAGATCTATCTTCACTTACTGATTAAGAACTCTGGCCATGCAAAAATTCGAGCACT 780
Db 1123 AAAGATCTATCTTCACTTACTGATTAAGAACTCTGGCCATGCAAAAATTCGAGCACT 1182
QY 781 AGCTGAAGAGTCTTAGCAGCAATCAGAACTGATTTGCTTGGAGCAAAAGAAAGA 840
Db 1183 AGCTGAAGAGTCTTAGCAGCAATCAGAACTGATTTGCTTGGAGCAAAAGAAAGA 1242
QY 841 ACTTGAAGGTACACAAAATTTAGAGAAGCTAAAGAAATTTGGATTAAGAAAGCTAT 900
Db 1243 ACTTGAAGGTACACAAAATTTAGAGAAGCTAAAGAAATTTGGATTAAGAAAGCTAT 1302
QY 901 CAGGCCAACATTTCTATTTGCTGCGCTTCTTATTTGATCTATGATCATATGCTGGC 960
Db 1303 TACAGCAATATTTCTATTTAGTGTGCTGCTTCTGCTGATCTATGATCTTATGCTGGC 1362
QY 961 TTTCTGATGAGGACCTCTTGTGCTCTCCAGTAAATTTCTATTTGAGCAAGTACTAC 1020
Db 1363 CTTTGTGATGAGGACCTCTTGTGCTCTCCAGTAAATTTCTATTTGAGCAAGTACTAC 1422
QY 1021 TGTCTTCTTCTTGTATTAATTTGAGGCTTTTATGATTTGAGCAAGCAATCCCAACATTTGA 1080
Db 1423 TGTATTTCTTCTGATTAATTTGAGGCTTTTATGATTTGAGCAAGCAATCCCAACATTTGA 1482
QY 1081 ASCATTTGCAAAAGCAAGAGAGAGAGCTTATGAATCTTCAAGATTAATGACAAATTAAC 1140
Db 1483 ASCATTTGCAAAAGCAAGAGAGAGAGCTTATGAATCTTCAAGATTAATGACAAATTAAC 1542
QY 1141 AAGCATTTACAGTATTTGAGAGTGGAGCAATTAACAGATTAATTAAGGAAATTTGGA 1200
Db 1543 AAGTATTTACAGTATTTGAGAGTGGAGCAATTAACAGATTAATTAAGGAAATTTGGA 1602
QY 1201 ATTCAAAAATGTCATTTCACTTACCTTCTCGAAAAGAAATTTAAGATCTTAAAGGCTCT 1260
Db 1603 ATTCAAAAATGTCATTTCACTTACCTTCTCGAAAAGAAATTTAAGATCTTAAAGGCTCT 1662
QY 1261 CAACCTGAAGGTTTCAGAGTGGGAGAGACAGTGGGCTGGTTGGGAACAGTGGCTGCCGGA 1320
Db 1663 GAACCTGAAGGTTTCAGAGTGGGAGAGACAGTGGGCTGGTTGGGAACAGTGGCTGGGGA 1722
QY 1321 GAGCAGCAGCGTGTGATGATGAGAGGCTTATGACCCACAGATGAGGCTGTGTAT 1380
Db 1723 GAGCAGCAGCGTGTGATGATGAGAGGCTTATGACCCACAGATGAGGCTGTGTAT 1782
QY 1381 TGATGACAGAGCATTTAGAGCCATAATGTAAGGCACTCTGGGAAATTTACTGGTGTGT 1440
Db 1783 TGATGACAGAGCATTTAGAGCCATAATGTAAGGCACTCTGGGAAATTTACTGGTGTGT 1842
QY 1441 GAGTCAGAGGCTGTGTGTGTCACACAGATGAGTGAATAATTCGCTATGGCCGGA 1500
Db 1843 GAGTCAGAGGCTGTGTGTGTCACACAGATGAGTGAATAATTCGCTATGGCCGGA 1902
QY 1501 AAATGTACCATGATGAGATTTGAGAAAGCTTTAAGAAAGCAATGGCTATGATTTAT 1560
Db 1903 AAATGTACCATGATGAGATTTGAGAAAGCTTTAAGAAAGCAATGGCTATGATTTAT 1962
QY 1561 CATGAAGCACTAATTAATTTGACACTCTGTGTTGAGAGAGAGGCGCCAGCTGAGTGG 1620

Db 1963 CATGAAGCACTAATTAATTTGACACTCTGTGTTGAGAGAGAGGCGCCAGTTGAGTGG 2022
QY 1621 TGGACAGAAACAGAGATGCGCATTTGCTGCGGCTGTTGCGCAACCCCAAGATTTCTG 1680
Db 2023 TGGGCAAGAGAGAGATGCGCATTTGCGCAAGTGGCTGTTGCGCAACCCCAAGATTTCTG 2082
QY 1681 GCTGATGAGCAAGCTGATGCTGAGCACTGGAAGTGAAGCAAGTGGTTCAGGTGGCCCT 1740
Db 2083 GCTGATGAGCAAGCTGATGCTGAGCACTGGAAGTGAAGCAAGTGGTTCAGGTGGCCCT 2142
QY 1741 GGATTAAGGCCAGAAAAGGCGGACTACATTTGATAGTATGATGATGATGATGATGATG 1800
Db 2143 GGATTAAGGCCAGAAAAGGCGGACTACATTTGATAGTATGATGATGATGATGATGATG 2202
QY 1801 TTAATGCGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1860
Db 2203 TTAATGCGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 2262
QY 1861 TGAATCTCATGAAGAGAGAGGCAATTTACTTCAAACTTTGACAAATGCAAGAGCAAGAGAA 1920
Db 2263 TGAATCTCATGAAGAGAGAGGCAATTTACTTCAAACTTTGACAAATGCAAGAGCAAGAGAA 2322
QY 1921 TGAATTTGAGTTGAAATTTCCACTGCTGATATCCAAAGTGAAGTGAATGCTTGGAAAT 1980
Db 2323 TGAATTTGAGTTGAAATTTCCACTGCTGATATCCAAAGTGAAGTGAATGCTTGGAAAT 2382
QY 1981 GTCTCCAAAGATTTAGGCTGCTGATTTAATAAAGAGATTTCACTGCGAGAGTATCA 2040
Db 2383 GTCTCCAAAGATTTAGGCTGCTGATTTAATAAAGAGATTTCACTGCGAGAGTATCA 2442
QY 2041 TGCACACAGAGCCCAAGACAGAAAGCTTTGACAAAGAGAGTGAATGAAATGAC 2100
Db 2443 TGCACACAGAGCCCAAGACAGAAAGCTTTGACAAAGAGAGTGAATGAAATGAC 2502
QY 2101 TCCAGTTTCCTTCTGAGAGATTTGAAAGCTGATCACTCACTCACTCACTCACTCACT 2160
Db 2503 TCCAGTTTCCTTCTGAGAGATTTGAAAGCTGATCACTCACTCACTCACTCACTCACT 2562
QY 2161 TGGTATTTTGTGCTATTTAATAAGAGGCTTCAACAGAGATTTTCAATTAATTTTC 2220
Db 2563 TGGTATTTTGTGCTATTTAATAAGAGGCTTCAACAGAGATTTTCAATTAATTTTC 2622
QY 2221 AAGCATTTAGGAGTCTTCCGAGATGAGATCCCTGAACCAAAAGCAGAGATAGTA 2280
Db 2623 AAGCATTTAGGAGTCTTCCGAGATGAGATCCCTGAACCAAAAGCAGAGATAGTA 2682
QY 2281 CATGTTTCTGATTTGTTTCTAGTCTTGAATTTCTTATTAATTAATTTCTCTCA 2340
Db 2683 CATGTTTCTGATTTGTTTCTAGTCTTGAATTTCTTATTAATTAATTTCTCTCA 2742
QY 2341 GGGCTTCAATTTGGCAAAAGCTGGGAGATCCCTCACTAAAGGCGCTTCAATGATGTTT 2400
Db 2743 GGGCTTCAATTTGGCAAAAGCTGGGAGATCCCTCACTAAAGGCGCTTCAATGATGTTT 2802
QY 2401 CAGATTCATCTGAGACAGAGATGTCAGTGTGTTGATGACCTTAAAAAACACACAGAGAGC 2460
Db 2803 CAGATTCATCTGAGACAGAGATGTCAGTGTGTTGATGACCTTAAAAAACACACAGAGAGC 2862
QY 2461 ATTGAACACAGGCTTCCCAATGATGATGCGCTCAAGTTAAAGGGCTATAGTTCCAGGCT 2520
Db 2863 ATTGAACACAGGCTTCCCAATGATGATGCGCTCAAGTTAAAGGGCTATAGTTCCAGGCT 2922
QY 2521 TGTGTCAATTTACCAAAATATAGCAAAATCTTGGAGAGCAATTAATTAATCTTAATCTA 2580
Db 2923 TGTGTCAATTTACCAAAATATAGCAAAATCTTGGAGAGCAATTAATTAATCTTAATCTA 2982
QY 2581 TGGTTGCAATTTACCAATTTTACTTGTAGCAATTTGACCAATTTGCAATTTGCAATTTG 2640
Db 2983 TGGTTGCAATTTACCAATTTTACTTGTAGCAATTTGACCAATTTGCAATTTGCAATTTG 3042
QY 2641 TGTGTAATGAATAATGTTGCTGAGCAAGCACTGAAAGATTAAGAAAGAGTGAAGAGAGC 2700
Db 3043 TGTGTAATGAATAATGTTGCTGAGCAAGCACTGAAAGATTAAGAAAGAGTGAAGAGAGTGC 3102

QY	3781	GCATGGACACATCAACAGCTGCTGCTGCCAGAAAGCATCTATTTTCCATGGTCAGTCT	3840		
Db	4183	GCATGGACACCATCACCAGCTGCTGCTGCCAGAAAGCATCTATTTTCCATGGTCAGTCT	4242		
QY	3841	CCAGGCTGGAGCAAGCGCTAGTGAACCTGCTGCCATATGAGCTGTATTAATTTTTTAAT	3900		
Db	4243	CCAGGCTGGAGCAAGCGCGCCAGAACTCTGACTGTATGAGATGTATTAATACCTTTTAAT	4302		
QY	3901	ATTGTGTAAACATGATGCTTTAATCAAAAGTTAAAGGTGAGACCTTACTGGAAAACT	3960		
Db	4303	ATT-TGTTTGAATATGACATTTATTCAAAGTTAA-----AGCAAAACCTTACAGATT	4355		
QY	3961	ATGTAGAATCTACTGTTTAACTTTCTGTCGCAACTGACATGACATTCACCAAGTTCAAG	4020		
Db	4356	ATGAGAGATATCTGTTTAACTTTCTGTCGTCGCAACTGACATTCACCAAGTTCAAG	4403		
QY	4021	AGCTCTGATTTATTAATTAAGGAACCA--AAGAAACATTATCTATGTGGATTAAT	4078		
Db	4404	-----AGGACTTCTGATTAATTAAGGAACAGAGGAGACATCAACAGTGGAGAAAT	4458		
QY	4079	ACTGCTGTTAATGCTATTAATTAATTTATAGTAATTCGAAGTATTTGTTATTAAT	4138		
Db	4459	CATAGTTTAACTGATTAATTAATTTATTAACGAAATTTAAAGTATTTTAAACATTA	4518		
QY	4139	TGTAATTAATTTGTTATTAATTTT-----ATTGTAACTACTGCTTGTGTAAGAT	4190		
Db	4519	ATGTGTAATTTTGTATTAATTTTCCCATTTGCACTGACTGACTGCTGCTGTAAGAT	4578		
QY	4191	TATAGAAGTGTAAAGAAAGTACTG--AATGTTGATAAGTAAAGCTAGCTATTAATTAAC	4249		
Db	4579	TATAGAAGTGTAAAGAAAGTATTAAGTATTAAGTATTAAGTATTAAGTATTAAGT	4635		
QY	4250	ACTTTTATAT 4259			
Db	4636	ACTTTCATGT 4645			
RESULT 13					
ABL68592	ID ABL68592 standard; DNA; 4646 BP.				
AC	ABL68592;				
DT	15-MAY-2002 (first entry)				
DE	Kidney cancer related gene sequence SEQ ID NO:6929.				
XX					
XX	Human; cancer; colon; breast; ovary; oesophagus; kidney; thyroid;				
KW	stomach; lung; prostate; pancreas; carcinoma; antitumour; cancerous;				
KW	cytostatic; gene therapy; antineoplastic; Wilm's tumour; adenocarcinoma;				
KW	gene; ds.				
OS	Homo sapiens.				
XX	WO200194629-A2.				
PN	13-DEC-2001.				
XX					
PF	30-MAY-2001; 2001WO-US10838.				
XX					
PR	05-JUN-2000; 2000US-209473P.				
PR	05-JUN-2000; 2000US-209531P.				
PR	18-SEP-2000; 2000US-233133P.				
PR	18-SEP-2000; 2000US-233617P.				
PR	20-SEP-2000; 2000US-234009P.				
PR	20-SEP-2000; 2000US-234034P.				
PR	20-SEP-2000; 2000US-234052P.				
PR	22-SEP-2000; 2000US-234509P.				
PR	22-SEP-2000; 2000US-234567P.				
PR	25-SEP-2000; 2000US-234923P.				
PR	25-SEP-2000; 2000US-234924P.				
PR	25-SEP-2000; 2000US-235077P.				
PR	25-SEP-2000; 2000US-235082P.				

[illegible]

D6 1003 CTTTCAGTCAATGCGACAATTTTTCACGCGTTTATATAGTATGGATTATTAACGCGGTGGAA 1062

D7 661 GCTAACCCCTTGTGATTTTGGCCATCAGCCCTGTTGTGACTTTCAGCGCGCATCTGTGGC 720

D8 1063 GCTAACCCCTTGTGATTTTGGCCATCAGCCCTGTTGTGACTTTCAGCGCGCATCTGTGGC 720

QY	721	AAAGATCTATCTTCATTTTACTGATATAAGAACTCTTGCGCTTGCAAAAGCTGGAGCAGT	780
Db	1123	AAAGATCTATCTTCATTTTACTGATATAAGAACTCTTGCGCTTGCAAAAGCTGGAGCAGT	1182
QY	781	ACCTGAGAGAGCTTAGAGCAATCAGAACTGTGATGTGCTTTGGAGACAAAAGAAAGA	840
Db	1183	ACCTGAGAGAGCTTAGAGCAATCAGAACTGTGATGTGCTTTGGAGACAAAAGAAAGA	1242
QY	841	ACTTGAAGAGTCAACAAATTTTAAAGAAAGCTTAAAGAAATTTGGGATTTAAAGAAAGCTAT	900
Db	1243	ACTTGAAGAGTCAACAAATTTTAAAGAAAGCTTAAAGAAATTTGGGATTTAAAGAAAGCTAT	1302
QY	901	CACGCGCAACATTTTCATTTGTGTGCGCTTCTTATGTATGTATGCATCATATGCTGTGC	960
Db	1303	TACAGCCATATTTTCATATAGTGTCGCTTTTCCCTGCTGATCATCATCTTATGCTGTGC	1362
QY	961	TTTCGTGATATGGAGACCTTCCTGTGGTCTCTCCAGTAAATTTCTATTTGGAGAAATATATTC	1000

[illegible]

QY	1081	1483	1441	1543	1542	1240
AGCATTTTGGCAAGGAGAGAGAGCATTTATGAATTTTAAATATATTTGACATTAAC	AGCATTTTGGCAAGGAGAGAGCATTTATGAATTTTAAATATATTTGACATTAAC	AGCATTTTGGCAAGGAGAGAGCATTTATGAATTTTAAATATATTTGACATTAAC	AGCATTTTGGCAAGGAGAGCATTTATGAATTTTAAATATATTTGACATTAAC	AGCATTTTGGCAAGGAGAGCATTTATGAATTTTAAATATATTTGACATTAAC	AGCATTTTGGCAAGGAGAGCATTTATGAATTTTAAATATATTTGACATTAAC	AGCATTTTGGCAAGGAGAGCATTTATGAATTTTAAATATATTTGACATTAAC

Oy	1201	ATTTCAAAAATGTCATCTCAGTTACCTTCTCGAAAAAGATTAAATCTTAAAGAGGTCT	1260
Oy	1202		1260
Db	1603	ATTTCAGAAATGTCTACTCTTACGTTACCCATCTCGAAAAAGATTAAAGATCTTGAAGAGGCTCT	1662
Oy	1261	CAACCTGAAGTTTCAGAGTGGGGCAGACAGTGGGGCGGTGGGAAACAGTGGCTGGCGGAA	1320
Oy	1262		1320
Db	1663	GAACCTGAAGGGGCGAGAGTGGGGCAGACAGTGGGGCGGTGGGAAACAGTGGCTGGCGGAA	1722
Oy	1321	GAGCAGACCTGGCGAGCTGATGCAGAGGGCTTATGACCCACAGATGGCAGGGGTGCTAT	1380
Oy	1723	GAGCAGACAGTCCAGCTGATGCGAGAGGCTCTATGACCCACAGGGGGATGGTCACTGTG	1782
Oy	1381	TGATGACAGAGCAATTAGGACCATTAATGTAAAGCACTCTTCGGGAAATTTACTGTGTGT	1440
Oy	1783	TGATGACAGAGCAATTAGGACCATTAATGTAAAGTTTCTACGGGAAATCATTTGGTGTGTG	1842
Db	1441	GAGTCAGAGACCTGTGTGTGTGTCACACAGATAGCTGAAAAACATTCGTTGGGCGCGA	1500
Oy	1442		1500
Db	1843	GAGTCAGAGACCTGTGTGTGTGTCACACAGATAGCTGAAAAACATTCGCTAGAGCGCGGA	1902
Oy	1501	AAATCCACCATGATGATGAGATTGAGAAAGCTGTGTAAAGAACCCATGCTATGATTTAT	1560
Oy	1903	AAATCTCACATGATGATGATGAGAAAGCTGTCAAGAAAGCCAAATGCTATGATTTAT	1962
Db	1561	CATGAACCTACTTAATTAATTTGACACTCTGTGTGGAGAGAGGGGCCACCTAGTGG	1620
Oy	1963	CATGAACCTGCTCTCTAATTTTGCACACCTGTGTGGAGAGAGGGGCCACCTAGTGG	2022
Db	1621	TGGAGAGAAACAGAGATGCCATTTGCTCGGGCCCTGTGTTGCACACCCCAAGTTCCT	1680
Oy	2023	TGGGACAGAGAGAGAGATGCCCATTTGCACGTCCCTGTGTTGCACACCCCAAGTTCCT	2082
Db	1681	GCTGATAGAGCAACCTGACCTCTGAGCACTGAAGGTGAACCACTGGTTAGGTGGCCT	1740
Oy	2083	GCTGATAGAGCCACGTCACCTCTTGGACACAGAAAGCAACCACTGGTTAGGTGGCCT	2142
Db	1741	GGATTAAGGCGAGAAAGGCCGGACTACCAATGTGATGACTATGCTTGTGCTCAAGTTTG	1800
Oy	2143	GGATTAAGGCGAGAAAGGTCGAGACCAATGTGATGACTATGCTTGTGCTCAAGTTTG	2202
Db	1801	TAAATCCCATGTCTATGCTGTGTTTGTGATGAGCACTGTGCGAAGAAAGAAATCATGA	1860
Oy	2203	TAAATGTCAGCTATGCTGTGTTTGTGATGAGCACTGTGCGAAGAAAGAAATCATGA	2262
Db	1861	TGAACCTCAGAAAGAGAGGCACTTACTTCAAACTTGTGCACAAATGCACAGAGAGAA	1920
Oy	2263	TGAACCTCAGAAAGAGAGGCACTTACTTCAAACTTGTGCACAAATGCACAGAGAGAA	2322
Db	1921	TGAATTTGAGTTAGAAATGCGACCTGGGTAATCCAAAACTGAAATGATGCTTGGAAAT	1980
Oy	2323	TGAATTTGAGTTAGAAATGCGACCTGATGAAATCCAAAGAGAAATTTGATGCTTGGAAAT	2382
Db	1981	GTCCTCAAAAGATTTCAGGCTCCAGTTTATATATAAAAGAAATCACTCCGAGAGATACA	2040
Oy	2383	GTCCTCAAAAGATTTCAGAGTCCAGCTTATATATAAGAAAGATCACTCGTAGAGATGTCG	2442
Db	2041	TGCACCAAGAGGCGAGAGAGAAACCTGTGTCACAAAGAGAGACTGATAGATGTRCC	2100
Oy	2443	TGCATCAGAGCCAGAGACAGAAACCTAGTACCAAGAGAGGCTGTGGAGAAAGTATACC	2502
Db	2101	TCCAGTTTCCCTTGGAGATTCCTGAACTGAATCAACAGTAAAGGCTTATTTGTGGT	2160
Oy	2503	TCCAGTTTCCCTTTGGAGGATTAATAGATCAATTAATTTACGTAAATGSCCTTATTTGCTGT	2562
Db	2161	TGATATATTTTGTGCTATTTAAAGCGAGGCCCTGCACACAGATTTTCAATATATTTTC	2220
Oy	2563	TGATATATTTTGTGCCATTTAATATGAGAGGCTGCACACAGCAATTTGCAATATATTTTC	2622
Db	2221	AAGATATATAGGATCTTTTACCAGATAGGATGCTGTAACCAAAAGCAGAGATATAGTA	2280
Oy	2623	AAGATATATAGGATCTTTTACAAAGATATATATCTTGAACCAAAAGCAGAGATATAGTA	2682

QY	2261	CATGTTTCTGATATGTTTCTTACGTCCTGGAAATTAATTTCTTTATTTATACATTTTTCCTCA	2340
Db	2683	CTTGTTTACATATTTGTTTTCAGCCCTGGAAATTTCTTTATTTATACATTTTTCCTTCA	2742
QY	2341	GGGCTTCATTTGGCAAGCTGGGGGATTCACATAAGCGGCTCGATATGATGTTT	2400
Db	2743	GGGTTTCATTTGGCAAGCTGGAGAGATCTCACAGCGGCTCGATATGATGTTT	2802
QY	2401	CAGATCCATGCGGAGACAGATGTCAGCTGGTTGATGACCTTAATAACACATGAGAC	2460
Db	2803	CCGATCCATGCTCAGACAGATGATGATGGTTGATGACCTTAATAACACATGAGAC	2862
QY	2461	ATTGACACACGAGCTTGGCATATATGGGGGTCAGTTAAAGGGGCTATATGTTCCAGCT	2520
Db	2863	ATTGATCCACGAGCTGGCAATGATGCTCTCAAGTTAAAGGGGCTATATGTTCCAGCT	2922
QY	2521	TGCTGCATTAACCCAGAAATATAGCAAACTTGGACAGAGCATTTATATCCTTAATCTA	2580
Db	2923	TGCTGTAATATCCAGAAATATAGCAAACTTGGACAGAGCAATTTATATCCTTCACTA	2982
QY	2581	TGGTTGGCAATTACACTTTTACTCTTGGCAATGTATCCATCTTCATATAGCAGAGT	2640
Db	2983	TGTTGGCAACTAACCTCTTACTCTTGGCAATGTATCCATCTTCATATAGCAGAGT	3042
QY	2641	TGTTGAAATGAAATATGTTCTGAGACAGACATGAAGATTAAGAAAGCTAGAAAGAGC	2700
Db	3043	TGTTGAAATGAAATATGTTCTGAGAACACATGAAGATTAAGAAAGCAATGAAGAGTGC	3102
QY	2701	TGGGAAAGATTGTAACAGAAAGCCATGCAAAATCTCCGAAGCTGTTGTTTGGACTCGGA	2760
Db	3103	TGGGAAAGATTGTAACAGAAATTAAGAAATCTCCGAAGCTGTTTGGACTCGAGGA	3162
QY	2761	GCAGAAATTTGAATATCATGTATGTCAGAGAGTTGGCAATGATCCATTCAGAAATCTTTAG	2820
Db	3163	GCAGAAATTTGAATATGATGATGCTCAGAGTTGGCAGTACCATATCAGAAATCTTTAG	3222
QY	2821	GAAAGCACACATCTTGGGGTCTCATTTTCTATCCACCGAGCAATGATATTTTCTTA	2880
Db	3223	GAAAGCACACATCTTGGAAATTCATTTTCTTACCTTCCACCGAGCAATGATATTTTCTTA	3282
QY	2881	TGCGGGCTGTTCCGGTTTGGTGGCTACTTGGTGGCAATGATTCATGACTTTCAGGA	2940
Db	3283	TGCTGATGTTTCCGGTTTGGAGGCTACTGTTGGTGGCAATGATTCATGACTTTCAGGA	3342
QY	2941	TGTTCTTTGGTATTTCTCAGCTATATGTCCTTGGTGGCCATGGCAGTGGGAGGTCAGTTC	3000
Db	3343	TGTTCTGTTAGATATTTTCAGCTGTTGTCCTTGGTGGCCATGGCAGTGGGGAATGTCAGTTC	3402
QY	3001	ATTGTCCTCGATATGTCCAAGCCAAAGATATCAGACCCAGCTCATATATCATTTGA	3060
Db	3403	ATTGTCCTCGATATGTCCAAAAGCAAAATATCAGACCCAGCTCATATGATCATTTGA	3462
QY	3061	AAAAAGCCCTGTATTTGACAGCTACAGCCCTCAGGCGCTCAAGCCAAATACGTTTGAAG	3120
Db	3463	AAAAAGCCCTGTATTTGATGACAGCTACAGCCGAAAGGCTTAATGCCGAACATTTGAAG	3522
QY	3121	AAATGTGACATTTAATGAGAGTCTGTTCACTATCCACTGACACAGATCCCGTGTCT	3180
Db	3523	AAATGTGACATTTGTTGTAAGTTGTATTAACATATCCACCGACCGGACATATCCAGTGTCT	3582
QY	3181	CCAGGGGCTGAGCCTTGAGGTAAAGAGGCCAGAGCGCTGGCCCTCGATGATAGACATGG	3240
Db	3583	TTCAGGGACTGAGCCTTGAGGTGAAGAGGCCAGAGCGCTGTGTGTGGGACACAGATGG	3642
QY	3241	CTTGGGGAAGACAGAGTTGTTTCAGCTCCTCATAAGAGCTTCTATATGACCCCTTGCTGTTC	3300
Db	3643	CTTGGGGAAGACAGACAGTGTCTCAGGTCTCTGGAGGGGTCTCTACGACCCCTTGGAGGGAA	3702
QY	3301	AGTGCATATTTGATGCAAGAGATTAAGACACTGATATGTCAGTGTCCGAGCACACT	3360
Db	3703	AGTGCATCTTGATGCAAGAAATAAGAGCACTGATATGTCAGTGTCTCCGAGCACACT	3762
QY	3361	GGGATGCTGTCTCAGGAGGCCATCTGTTTGACTCAGACATTTGCCGAGACATTTGCTTA	3420


```

Db 3763 GGGGATCGTGTCCAGAGAGCCCATCTGTTGACTGCAGATTTGCGAATAATTGCTTA 3822
OY 3421 TGGAGCAACAGCCGGGTGCTATCAGATGAAGATTATGACGAGCCAGCCAGAGCCAA 3480
Db 3823 TGGAGCAACAGCCGGGTGCTATCAGATGAAGATTATGACGAGCCAGCCAGAGCCAA 3882
OY 3481 CATACACACTTCATCAGAGACTCCCTGAGAAATCAACAGACAGATGAGACAAAG 3540
Db 3883 CATACAGCTTCATCAGAGACTCCCTGAGAAATCAACAGACAGATGAGACAAAG 3942
OY 3541 AACCAGCTCTGTTGGGCGACAGAACGCGATTGCCATAGCTCGCTGTTTAAACA 3600
Db 3943 AATCAGCTCTGTTGGGCGACAGAACGCGATTGCCATAGCTCGCTGTTTAAACA 4002
OY 3601 GCCTCATATTTTCTTTTGGATGAGAGCTCATCAGCTGCTGATACAGAAAGT 3660
Db 4003 GCCTCATATTTTCTTTTGGATGAGAGCTCATCAGCTGCTGATACAGAAAGT 4062
OY 3661 TGTCCAAAGAGCCCTGAGCAAAAGCCAGAGAGGCGCGACCTGATGATCGCCACCG 3720
Db 4063 TGTCCAAAGAGCCCTGAGCAAAAGCCAGAGAGGCGCGACCTGATGATCGCTCACCG 4122
OY 3721 CTTGTCCACCATCCAGATGAGATGATGATGATGATGATGATGATGATGATGATGAT 3780
Db 4123 CTTGTCCACCATCCAGATGAGATGATGATGATGATGATGATGATGATGATGATGAT 4182
OY 3781 GCATGCGACATCAACAGCTGCTGCGCCAGAAAGCATTTTTCATGCTGCTGCTGCT 3840
Db 4183 GCATGCGACATCAACAGCTGCTGCGCCAGAAAGCATTTTTCATGCTGCTGCTGCT 4242
OY 3841 CCAGGCTGGAGAGAGCCCTGAGTACGCTGAGTACGCTGAGTACGCTGAGTACGCTGAG 3900
Db 4243 CCAGGCTGGAGAGAGCCCTGAGTACGCTGAGTACGCTGAGTACGCTGAGTACGCTGAG 4302
OY 3901 ATTGTGTTAAACATGAGCATTTATCAAAAGTAAAGAGGAGCACTTACGGAAGAACT 3960
Db 4303 ATT--TGTTAGATATGACATTTTATCAAGTTAA--AGCAACGCTTACGGAAT 4355
OY 3961 ATGTAGAACTACCTGTTTAACTTCTGCTGCACTGAGATCAATTCACCAAGTTCAG 4020
Db 4356 ATGTAGAACTACCTGTTTAACTTCTGCTGCACTGAGATCAATTCACCAAGTTCAG 4403
OY 4021 AGTCTTCAATTTTATTAATTAAGAACCA--AAGAACTTATCTGATGATGATGAT 4078
Db 4404 ----AGAGACTTCGTAAATTAAGAACAGAGTGAAGACATCAAGTGAAGAAAT 4458
OY 4079 ACTGCTTATATGCACTTAAATTAATTAAGAACTTATGAGTAATTCAGAGTATTTGTAAT 4138
Db 4459 CATAGTTTAACTGCAATTAATTAATTTATTAACAGATTAAGTAAAGTAAAGATAA 4518
OY 4139 TGTATATTTTGTATATATTTT-----ATTGTAAGTACTGCTTGGCTGAAGAT 4190
Db 4519 ATGTATATTTTGTATATATTTTCCCATTTGAGACTGATGAGTACTGCTGCTGAAGAT 4578
OY 4191 TATAGAAGTGTAAAGAAAGTACTG--AATGTTGAATTAAGTCTGATATATTAAGTAA 4249
Db 4579 TATAGAAGTGTAAAGAAAGTACTG--AATGTTGAATTAAGTCTGATATATTAAGTAA 4635
OY 4250 ACTTTTATAT 4259
Db 4636 ACTTTCATGT 4645

```

RESULT 14

ABL68880

ID ABL68880 standard; DNA; 4646 BP.

AC ABL68880;

DT 15-MAY-2002 (first entry)

DE Kidney cancer related gene sequence SPO ID NO:7217.

```

XX Human; Cancer; colon; breast; ovary; oesophagus; kidney; thyroid;
KW stomach; lung; prostate; pancreas; carcinoma; antitumour;
KW cytosolic; gene therapy; antineoplastic; Wilm's tumour; adenocarcinoma;
XX gene; ds.
OS Homo sapiens.
XX
XX WO200194629-A2.
XX
XX 13-DEC-2001.
XX
XX 30-MAY-2001; 2001WO-US10838.
XX
XX 05-JUN-2000; 2000US-209473P.
XX 05-JUN-2000; 2000US-209531P.
XX 18-SEP-2000; 2000US-233133P.
XX 18-SEP-2000; 2000US-233617P.
XX 20-SEP-2000; 2000US-234009P.
XX 20-SEP-2000; 2000US-234034P.
XX 22-SEP-2000; 2000US-234052P.
XX 22-SEP-2000; 2000US-234509P.
XX 22-SEP-2000; 2000US-234567P.
XX 25-SEP-2000; 2000US-234923P.
XX 25-SEP-2000; 2000US-234924P.
XX 25-SEP-2000; 2000US-235077P.
XX 25-SEP-2000; 2000US-235082P.
XX 25-SEP-2000; 2000US-235134P.
XX 25-SEP-2000; 2000US-235280P.
XX 26-SEP-2000; 2000US-235337P.
XX 26-SEP-2000; 2000US-235638P.
XX 27-SEP-2000; 2000US-235711P.
XX 27-SEP-2000; 2000US-235720P.
XX 27-SEP-2000; 2000US-235840P.
XX 27-SEP-2000; 2000US-235863P.
XX 28-SEP-2000; 2000US-236028P.
XX 28-SEP-2000; 2000US-236032P.
XX 28-SEP-2000; 2000US-236033P.
XX 28-SEP-2000; 2000US-236034P.
XX 28-SEP-2000; 2000US-236109P.
XX 28-SEP-2000; 2000US-236111P.
XX 29-SEP-2000; 2000US-236842P.
XX 29-SEP-2000; 2000US-236849P.
XX 02-OCT-2000; 2000US-237172P.
XX 02-OCT-2000; 2000US-237173P.
XX 02-OCT-2000; 2000US-237175P.
XX 02-OCT-2000; 2000US-237278P.
XX 02-OCT-2000; 2000US-237294P.
XX 02-OCT-2000; 2000US-237295P.
XX 02-OCT-2000; 2000US-237316P.
XX 03-OCT-2000; 2000US-237425P.
XX 03-OCT-2000; 2000US-237588P.
XX 03-OCT-2000; 2000US-237604P.
XX 03-OCT-2000; 2000US-237606P.
XX 03-OCT-2000; 2000US-237608P.
XX 01-NOV-2000; 2000US-244867P.
XX 01-NOV-2000; 2000US-245084P.
XX
XX (AVAL-) AVALON PHARM.
XX
XX Young PE, Augustus M, Carter KC, Ebner R, Endress G, Horrigan S;
XX Sopet DR, Weaver Z;
XX
XX WPI; 2002-188264/24.
XX
XX Screening for anti-neoplastic agent involves exposing cells to a
XX chemical agent to be tested for anti-neoplastic activity, and
XX determining a change in expression of a gene of a signature gene set
XX
XX Claim 1; SEQ ID 7217; 44pp; English.
XX
XX The present invention describes a method (M1) for screening for an
XX anti-neoplastic agent. The method involves exposing cells to a chemical
XX agent to be tested for anti-neoplastic activity, determining a change in

```

expression of at least one gene (1) of a signature gene set, where (1) comprises a sequence (S) selected from 8447 sequences (given in ABU61664 to ABU70110), or is at least 95% identical to (S), where a change in expression is indicative of anti-neoplastic activity. (1) has cytostatic activity and can be used in gene therapy. M1 can be used for screening an anti-neoplastic agent, and can be used for producing a product which is the data collected with respect to the anti-neoplastic agent as a result of M1, and the data is sufficient to convey the chemical structure and/or properties of the agent. M1 can be used in the treatment of cancer such as colon, breast, stomach, lung, thyroid, oesophageal, ovarian, kidney, prostate or pancreatic cancer, adenocarcinoma, carcinoma, clear cell cancer, infiltrating ductal cancer, infiltrating lobular cancer, squamous cell carcinoma, neuroendocrine carcinoma, papillary carcinoma and Wilms' tumour.

XX Sequence 4646 BP; 1371 A; 892 C; 1129 G; 1254 T; 0 other;

SQ

Query Match 77.7%; Score 3324.4; DB 24; Length 4646;

Best Local Similarity 87.6%; Pred. No. 0;

Matches 3740; Conservative 0; Mismatches 486; Indels 44; Gaps 8;

QY 1 GGAGCGCGAGTGGGATGATCTCTGAAGAGAGCCGTAAAGGAGTGCAGAGAAACTT 60

DB 409 GGAGCGCGAGTGGGATGATCTCTGAAGAGAGCCGTAAAGGAGTGCAGAGAAACTT 60

QY 61 CTGGAATAAGGCAAAAAAGTAAAAAATGAGAAAGAAAGAAAGAAACCACTGTAC 120

DB 469 CTTTAAAACTGACAAATTAAGTAAAAAGATTAAGAAAGAAAGAAACCACTGTAC 120

QY 121 CAGCTTGCATGTTTCGATCAATAGTGGCTTGAAGTTGATATGTTGGGGAC 180

DB 529 TGTATTTTCATGTTTCGATCAATAGTGGCTTGAAGTTGATATGTTGGGGAC 180

QY 181 AATGGCTGCATCATTCATGAGCTGCACCTCCATCATGATGCTGTTTGGAAACAT 240

DB 589 TTTGGCTGCATCATTCATGAGCTGCACCTCCATCATGATGCTGTTTGGAAACAT 240

QY 241 GAGCATAGCTTTCATCAATGACAGAAATTTCAAGAAACAAACTTTCCAGTTATATTA 300

DB 649 GACGATATCTTTCATCAATGACAGAAATTTCAAGAAATCTCATGTCA-----ACATCAC 702

QY 301 TGAATATTAAGCAATACATACACATTTTCATCATCAACATTCGAGAGAAATGACAC 360

DB 703 TAAATGAATGATATCATATACAGGATCTTCATGATATGAGAGAAATGACAC 360

QY 361 GATATGATATTAATCAATGAGTGCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 420

DB 763 GATATGATATTAATCAATGAGTGCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 420

QY 421 TTTATTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 480

DB 823 TTTATTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 480

QY 481 TGTATATGATGAGAGAGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 540

DB 883 TGTATATGATGAGAGAGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 540

QY 541 CCGGCTCAGACAGATGCTCCAAATCAATGAAGAAATGGGAGCAAAATTTGAAATGTT 600

DB 943 CCGGCTCAGACAGATGCTCCAAATCAATGAAGAAATTTGAAATTTGAAATGTT 600

QY 601 CTTTCACTCAATGAAGCAATTTTTCACCGTTTATGAGGTTTACAGCTGTTGAA 660

DB 1003 CTTTCACTCAATGAAGCAATTTTTCACCGTTTATGAGGTTTACAGCTGTTGAA 660

QY 661 GCTAACCTTGTGATTTGGCCATCAGCCCTTCTTGGATCTTCAAGCCCTGAGG 720

DB 1063 GCTAACCTTGTGATTTGGCCATCAGCCCTTCTTGGATCTTCAAGCCCTGAGG 720

QY 721 AAGATATCATCTTCAATTAAGAAAGCTTGGCTTGAAGAAAGCTTGAAGAGT 780

DB 1123 AAGATATCATCTTCAATTAAGAAAGCTTGGCTTGAAGAAAGCTTGAAGAGT 780

QY 781 AGCTGAAGAAGCTTTAGCAGCAATCAGACACTGATTCGCTTGGAGCAAAAGAAAGA 840

DB 1183 AGCTGAAGAAGCTTTAGCAGCAATCAGACACTGATTCGCTTGGAGCAAAAGAAAGA 840

QY 841 ACTTGAAGGTCACCAAAATTTGAAGAAGCTTAAGAAATTTGGATTAAGAAAGCT 900

DB 1243 ACTTGAAGGTCACCAAAATTTGAAGAAGCTTAAGAAATTTGGATTAAGAAAGCT 900

QY 901 CAGGAGCAACATTTGATGAGTGCCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 960

DB 1303 TACAGCCAAATTTGATAGTGCCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 960

QY 961 TTTCTGATGAGGAGCTCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1020

DB 1363 CTTCTGATGAGGAGCTCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1020

QY 1021 TGTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 1080

DB 1423 TGTATTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 1080

QY 1081 AGCATTTGCAAG 1140

DB 1483 AGCATTTGCAAG 1140

QY 1141 AAGCATTTGCAAG 1200

DB 1543 AAGCATTTGCAAG 1200

QY 1201 ATTCAAAATTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 1260

DB 1603 ATTCAAAATTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 1260

QY 1261 CAACCTGAAGTTGAG 1320

DB 1663 CAACCTGAAGTTGAG 1320

QY 1321 GAGCAG 1380

DB 1723 GAGCAG 1380

QY 1381 TGATGAG 1440

DB 1783 TGATGAG 1440

QY 1441 GAGTCAAG 1500

DB 1843 GAGTCAAG 1500

QY 1501 AAATGTCACCATGATGAGATGAGAAAGCTGTTAAGAAAGCTGTTAAGAAAGCTGTTAAG 1560

DB 1903 AAATGTCACCATGATGAGATGAGAAAGCTGTTAAGAAAGCTGTTAAGAAAGCTGTTAAG 1560

QY 1561 CATGAATCTGCTCATTAATTTGACATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1620

DB 1963 CATGAATCTGCTCATTAATTTGACATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1620

QY 1621 TGGACAGAAACAG 1680

DB 2023 TGGACAGAAACAG 1680

QY 1681 GCTGATGAG 1740

DB 2083 GCTGATGAG 1740

QY 1741 GGATTAAG 1800

DB 2143 GGATTAAG 1800

QY 1801 TAAATGCTGATGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1860

DB 2203 TAAATGCTGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1860

QY 1861 TGAATCATTAAGAAAG 1920

Db 2263 TGAAGTCAATGAAAGAGGACATTTACTTCAAACTTGTCACAAATGAGACACAGAGAAA 2322
QY 1921 TGAATTTAGTTAGAAAATGCGCACTGGTGAATCCAAAAAGTGAAGATGATCCCTTGGAAAT 1980
Db 2323 TGAATTTAAATTAAGAAAATGACGTGATGAATCCAAAAAGTGAATGATGATCCCTTGGAAAT 2382
QY 1981 GTCTCCAAAAGATTCAGGAGTCCAGTTTAATAAAAAGAGATGCAATCCGACAGATATACA 2040
Db 2383 GTCTTCAAAATGATTCAGATCCAGTCTTAATAAAAAGAGATGCAATCCGACAGATATACA 2442
QY 2041 TGCACCAAGGCCACAGACAGAAAGCTTGGTGAACAAAAAGAGATGCAATGATGATCC 2100
Db 2443 TGGATCCAAAGGCCACAGACAGAAAGCTTGGTGAACAAAAAGAGATGCAATGATGATCC 2502
QY 2101 TCCAGTTTCCCTTGGAGATTTCTGAAGCTTGAACCTCAATGATGATGATGATGATGATG 2160
Db 2503 TCCAGTTTCCCTTGGAGATTTATGAAGCTTGAATTTAATGATGATGATGATGATGATGATG 2562
QY 2161 TGGATATTTTGTGCTATTTAATAACGAGAGCCCTGCACACAGATTTTCAATATATTTTC 2220
Db 2563 TGGATATTTTGTGCTATTTAATAACGAGAGCCCTGCACACAGATTTTCAATATATTTTC 2622
QY 2221 AAGGATTAATAGGATCTTTAACCCGAGATGAGATCCTGAAACAAACAGACAAATAGTAA 2280
Db 2623 AAGGATTAATAGGAGTTTTCACAAAGATGATGATCCTGAAACAAACAGACAAATAGTAA 2682
QY 2281 CATGTTTCTGTATGTTTCTATGCTTGAATTTATTTCTTTTATATATTTTCCCTCCA 2340
Db 2683 CTGTTTCTACTATGTTTCTAGCCCTTGAAATTTTCTTTTATTTTCCCTCCA 2742
QY 2341 GGGCTTCACATTTGGCAAAAGCTGGGAGATCCTCAATAGCGGCTGGATACATGTTT 2400
Db 2743 GGGTTTCACATTTGGCAAAAGCTGGGAGATCCTCAATAGCGGCTGGATACATGTTT 2802
QY 2401 CAGATCCATGCTGAGACAGAGATGTCAGCTGTTGATGATGATGATGATGATGATGATGATG 2460
Db 2803 CAGATCCATGCTGAGACAGAGATGTCAGCTGTTGATGATGATGATGATGATGATGATGATG 2862
QY 2461 ATTGACACCGAGGCTTGGCAGATGATGCGGCTCAAGTTAAAGGGCTATAGTTCCAGGCT 2520
Db 2863 ATTGACACCGAGGCTTGGCAGATGATGCGGCTCAAGTTAAAGGGCTATAGTTCCAGGCT 2922
QY 2521 TGTGTCATTAACCCAGAAATATAGCAATCTTGGACAGGATATATATATATATATATATATAT 2580
Db 2923 TGTGTCATTAACCCAGAAATATAGCAATCTTGGACAGGATATATATATATATATATATATAT 2982
QY 2581 TGTGTCATTAACCACTTTTACTCTTGAATTTGATACCATCATTTGCAATAGCAGAGAT 2640
Db 2983 TGTGTCATTAACCACTTTTACTCTTGAATTTGATACCATCATTTGCAATAGCAGAGAT 3042
QY 2641 TGTGTAATGAAATGTTTGTCTGACAGACATGAAAGATGAAAGATGAAAGATGAAAGATGAA 2700
Db 3043 TGTGTAATGAAATGTTTGTCTGACAGACATGAAAGATGAAAGATGAAAGATGAAAGATGAA 3102
QY 2701 TGGGAAATGCTACAGAAAGCCATGCAAAACCTTCCGAACTGTTGTTCTTTTACTCGGGA 2760
Db 3103 TGGGAAATGCTACAGAAAGCCATGCAAAACCTTCCGAACTGTTGTTCTTTTACTCGGGA 3162
QY 2761 GCAGAAAGTTGAATATCATATGACAGAGTTGCAAGTACATACAGAAATCTTTGAG 2820
Db 3163 GCAGAAAGTTGAATATCATATGACAGAGTTGCAAGTACATACAGAAATCTTTGAG 3222
QY 2821 GAAGACACACATCTTGGGGGTCTCATTTTCTATCACCCAGGCAATGATGATTTTCTCTA 2880
Db 3223 GAAGACACACATCTTGGAAATATCATTTTCTTCCACCCAGGCAATGATGATTTTCTCTA 3282
QY 2881 TGTGTCATTTCCGTTTGGTGGTCTACTGTTGGCAATGATGATGATGATGATGATGATGATGATG 2940
Db 3283 TGTGTCATTTCCGTTTGGTGGTCTACTGTTGGCAATGATGATGATGATGATGATGATGATGATGATG 3342
QY 2941 TGTGTCATTTGGATTTCTAGCATTTGCTTGTGTCATGACATGAGGAGGAGGAGGAGGAGG 3000

Db 3343 TGTCTGTAGTATTTTTCAGCTGTTGTCTTTGGTGCATGAGCCGCTGGGCAATGACATTC 3402
QY 3001 ATTGCTCTGATCATATGCAAAAGCCAAAGTATGACAGCCAGCTCATCATATGATATGTA 3060
Db 3403 ATTGCTCTGATCATATGCAAAAGCCAAAGTATGACAGCCAGCTCATCATATGATATGTA 3462
QY 3061 AAAAGCCCTGATGATGACAGCTACAGCCCTCAGCGCTCAGCCCAATATGCTTGGAG 3120
Db 3463 AAAAGCCCTGATGATGACAGCTACAGCCAGGAGGCTCAATGCTCCAGACATTTGGAG 3522
QY 3121 AAATGTCATTTAATGAGGCTGTTCACTATCCACTGACACAGCATCCGCTGT 3180
Db 3523 AAATGTCATTTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 3582
QY 3181 CCAAGGCTGAGGCTGAGGCTGAGAGAGGCCAGACGCTGCGCTGTAGTATGACATG 3240
Db 3583 TCAAGGACTGAGGCTGAGGCTGAGAGAGGCCAGACGCTGCGCTGTAGTATGACATG 3642
QY 3241 CTGTGGGAAGACACATGTTTCAAGCTCCTAGAGGCTGTATGATGATGATGATGATGATG 3300
Db 3643 CTGTGGGAAGACACATGTTTCAAGCTCCTAGAGGCTGTATGATGATGATGATGATGATG 3702
QY 3301 AGTCTAATTTGATGCAAAAGATTAAGCACCTGAATGTCAGTGGCTCCGACACACT 3360
Db 3703 AGTCTAATTTGATGCAAAAGATTAAGCACCTGAATGTCAGTGGCTCCGACACACT 3762
QY 3361 GGGACGCTGCTCAGAGAGCCCATCTTGTGATGATGATGATGATGATGATGATGATGATG 3420
Db 3763 GGGACGCTGCTCAGAGAGCCCATCTTGTGATGATGATGATGATGATGATGATGATGATG 3822
QY 3421 TGGAGCAACAGCGGGTGTATCATGATGAGAGATTAATGACAGGACCCCAAGAGGCA 3480
Db 3823 TGGAGCAACAGCGGGTGTATCATGATGAGAGATTAATGACAGGACCCCAAGAGGCA 3882
QY 3481 CATACACCACTTCAATCGAGACATCCCTGAGAAATTAACAACAGATGAGAGCAAAAG 3540
Db 3883 CATACACCACTTCAATCGAGATCTGCTTAATTAATTAATGAGACATTAAGAGACAAAG 3942
QY 3541 AACCCAGCTCTGCTGAGCAAAAGCCAGATTTGCAATGCTGCGCTGTGATGATGATGATG 3600
Db 3943 AACCCAGCTCTGCTGAGCAAAAGCCAGATTTGCAATGCTGCGCTGTGATGATGATGATG 4002
QY 3601 GCTCATATTTTGTGTTTGGATGAGTACATCAGCTCTGATACAGAAAGTAAAGT 3660
Db 4003 GCTCATATTTTGTGTTTGGATGAGTACATCAGCTCTGATACAGAAAGTAAAGT 4062
QY 3661 TGTCCAAAGAGCCCTGAGCAAAAGCCAGAGAGGCGCACCTGATGATGATGATGATGATG 3720
Db 4063 TGTCCAAAGAGCCCTGAGCAAAAGCCAGAGAGGCGCACCTGATGATGATGATGATGATG 4122
QY 3721 CTGTCCACATCCAGATCCAGATTTAATAGTGGTGTTCAGAAATGCAAAAGTCAAG 3780
Db 4123 CTGTCCACATCCAGATCCAGATTTAATAGTGGTGTTCAGAAATGCAAAAGTCAAG 4182
QY 3781 GCATGACACATCAGCAGCTGCTGACAGAAAGGCAATGATTTTCAATGATGATGATG 3840
Db 4183 GCATGACACATCAGCAGCTGCTGACAGAAAGGCAATGATTTTCAATGATGATGATGATG 4242
QY 3841 CCAAGGCTGAGCAAAAGCCCTAGTGAACCTGTGACATATGATGATGATGATGATGATGATG 3900
Db 4243 CCAAGGCTGAGCAAAAGCCCTAGTGAACCTGTGACATATGATGATGATGATGATGATGATG 4302
QY 3901 ATTTGTTTAAACATGAGCATTTAATCAAAAGTTAAAGTGAAGCACTTACGAAAACT 3960
Db 4303 ATTTGTTTAAACATGAGCATTTAATCAAAAGTTAAAGTGAAGCACTTACGAAAACT 4355
QY 3961 ATTTGTTTAAACATGAGCATTTTAACTTTCTGTCAGTGAAGATTCACAAAGTTCAAG 4020
Db 4356 ATTTGTTTAAACATGAGCATTTTAACTTTCTGTCAGTGAAGATTCACAAAGTTCAAG 4403
QY 4021 AGTCTTCAAGATTTAATTAAGAGCAAA - AAGAAATATATGATGATGATGATGATGATGATG 4078
Db 4404 -----AGAGCTGTGTATTAAGAGAACAGATGAGAGACATCATCAAGTGAAGAAAT 4458

	Query Match	77.6%	Score 3322.6;	DB 22;	Length 4349;	
	Best Local Similarity	87.5%;	Pred. No. 0;	Mismatches 489;	Indels 44;	Gaps
	Matches 3740;	Conservative	0;			
QY	1	GGAGCGCAGGTCCGGAATGCATCTCTGAAGAGGCCGCTAAAGGGAGTGCAGAGAAGACTT	60			
Ddb	110	GGAGCGCAGGTCCGGAATGCATCTTTGAAGGACCCGCAATGAGAGAGCAAAAGAA	169			
QY	61	CTGAAAATTGGGCAAAAAAAGTAATAAAAAATGAGAGAAAAAGAAACAACCTGTCAG	120			
Ddb	170	CTTTTTTAACTGAACCATTAAGTAGAAAAAGATMAAGAAAAAGAACCAACTGTGAC	229			
QY	121	CACGTTTGCAATGTTCCCTATTCCAATTTGGCTTGATAGGTGTATATGTTGGTGGGAC	180			
Ddb	230	TGTATTTCGAAGTTTCCCTATTCTCAAATTTGGCTTGACAAAGTTATATGTTGGTGGGAC	289			
QY	181	AATGGCTGCATCATCATCATGAGCTGCACCTCCCTCATGATGCTGTTTTGGAAACT	240			
Ddb	290	TTTGGCTGCATCATCATCATGGGGCTGGCACTCTCTCATGATGCTGTGTTGGAGAAAT	349			
QY	241	GACAGATACCTTTGCCAATTCGAGAAATTTCAAGAAACAAACTTTCCACTTATATTA	300			
Ddb	350	GACAGATACCTTTGCCAATTCGAGAAATTTAGAAATCTGATGTCA-----AACATCAC	403			
QY	301	TGAAATATTAGCAACAAATACAAACATTTTCATCAACATTCGAGAGGAAATGACCAC	360			
Ddb	404	TAAATAGAGATATCAATGATACAGGGTCTTCAATGATCTGGAGAGAACTATCACAG	463			
QY	361	GTAATGCTTATTTACATGTGGATCGCTCTGGCGTGGCGTGGCTGTTACATCAAGT	420			
Ddb	464	GTAATGCTTATTTACATGTGGATCGCTCTGGCGTGGCGTGGCTGTTACATCAAGT	523			
QY	421	TTCATTTGGTGGCTGGCAGCAGAGAGACAGATATCTCAAAATTAAGAAAAATTTTTTCA	480			
Ddb	524	TTTCAATTTGGTGGCTGGCAGCAGTGAAGACAAATATCAAAATTAAGAAAAATTTTTTCA	583			
QY	481	TGCTATCATGCGACAGAGATGGCTGGTGTGACGTGACATGACGTGGGAGCTTAACAC	540			
Ddb	584	TGCTATCATGCGACAGAGATGGCTGGTGTGATGTGACAGAGTGTGGGAGCTTAACAC	643			
QY	541	CCGGCTCAGACAGCATGTCTCCAAATCAATGAAAGAAATTTGGCAGCAAAATTTGCAATGT	600			
Ddb	644	CCGACTTACAGATATATGTCTCTAAGATTAATGAGATTTATGGTGACAAAAATTTGCAATGT	703			
QY	601	CTTTCATCAATAGCAACATTTTACACCGCTTTATATGTTGGGTTTACAGTGTGGAA	660			
Ddb	704	CTTTCATCAATAGCAACATTTTACCTGGGTTTATATGATGAATTAACAAGTGTGGAA	763			
QY	661	GCTAACCTTGATTTTGGCCATCAGCCCTGTCTTGAATTTACAGCCGCCATCTGGGC	720			
Ddb	764	GCTAACCTTGATTTTGGCCATCAGCTCGTCTTGGAGCTGCACCTCTCTGTGGGC	823			
QY	721	AAAGATACTATCTTACTTACTGATTAAGAACTCTGTGGCTATGCAAAACCTGGAGAGT	780			
Ddb	824	AAAGATACTATCTTACTTACTGATTAAGAACTCTGTGGCTATGCAAAACCTGGAGAGT	883			
QY	781	ACCTGAAGAGCTTTAGCAGCACTCAGAACTGTGATTCCTTTGGAGACAAAAAGAAAGA	840			
Ddb	884	ACCTGAAGAGCTTTGGCAGCAATTAAGAACTGTGATTCATTTGGAGACAAAAAGAAAGA	943			
QY	841	ACTTGAAGAGTTCACAAAAATTTAGAAAGACTTAAGAGATTTGGATTAAGAAAGCAT	900			
Ddb	944	ACTTGAAGAGTTCACAAAAATTTAGAAAGACTTAAAGAAATTTGGATTAAGAAAGCAT	1003			
QY	901	CACGGCCAACTTTCTATGTGGTGGCGCTTCTTATGATCTATGCATATATGCTCTGGC	960			
Ddb	1004	TACAGCCAAATTTCTATGTGGTGGCTTCTTCTGCTGATCTATATGATCTTATGTCTGGC	1063			
QY	961	TTTTTGATGGAGACTCTGTGGTCTCTCCAGTGAATTTCTATTTGACAAAGTACTCAC	1020			
Ddb	1064	CTTCTGTATGGAGCAACCTGTGGTCTCTCCAGGGAATTTCTATTTGGCAAGTACTCAC	1123			

OY	1021	TCGCTCTCTTTCGTATTAATTGGGGCTTTTGGTATTTGGACAGGACATCCCAACCATTTGA	1080
Db	1124	TGTATTTCTTTTCGTATTAATTGGGGCTTTTGGTATTTGGACAGGACATCCCAACCATTTGA	1183
OY	1081	AGCATTTGGCAAGGAGGAGGAGACACCTTATCAAAATCTTCAAGATTAATTGGCATTAACC	1140
Db	1184	AGCATTTGGCAAAATGCAAGAGGAGGAGACCTTATTAATCTTCAAGATTAATTGGATTAAGCC	1243
OY	1141	AAGCATTGACAGCTATTTGCAAGAGTGGACATTAACACAGATATAATTAAGGAAATTTTGA	1200
Db	1244	AAGTATTTGACAGCTATTTGCAAGAGTGGGACAAACACAGATATAATTAAGGAAATTTTGA	1303
OY	1201	ATTCAAAATATGTTACCTCACTGTTACCTTCTCGAAAAGAGTTAAGATCTTTAAAGGCTCT	1260
Db	1304	ATTCGAAAATGTTACCTCACTGATTACCCATCTCGAAAAGAGTTAAGATCTTTGAAGGGCTCT	1363
OY	1261	CAACCTTAAGGTTCAAGTGGGGGACAGACATGGGCGGTGGGGAACATGGGCTCGGGGAA	1320
Db	1364	GAACCTTAAGGTTCAAGTGGGAGGAGGAGACATGGGCGGTGGGGAACATGGGCTCTGGGAA	1423
OY	1321	GAGCAGACACGCTGAGCTGATGTCAGAGGCTCTATGACCCACACAGATGGCATGCTGTAT	1380
Db	1424	GAGCACAACAGTCAAGCTGATGTCAGAGGCTCTATGACCCACACAGAGGGGATGTCAGAGT	1483
OY	1381	TGATGACACAGCATTTAGACCCATAATATTAAGCATCTTGCGGAAATTACTGGTGTGCT	1440
Db	1484	TGATGGACAGGATATTGAGCCATAATATTAAGGTTTCTACGGGAAATCAATTTGCTGTGCT	1543
OY	1441	GAGTCAGAGCGCTGCTGTTGTTGGCCACACAGATAGCTGAAACATTTGCTTATGGCCGGA	1500
Db	1544	GAGTCAGGAACTGTATTTGTTTCCACACAGATAGCTGAAACATTTGCTTATGGCCGCTGA	1603
OY	1501	AAATGTCAACATGATGATTTAGAAAGCTGTTAAGGAAGCCATGGCTATGATTTTAT	1560
Db	1604	AAATGTCAACATGATGATTTAGAAAGCTGTCAAGGAAGCCATGGCTATGATTTAT	1663
OY	1561	CATGAACCTACTTATTAATTTACACTCTGTTGGAGAGAGAGGGGCCACAGTGAAGTG	1620
Db	1664	CATGAACCTGCTCTATTAATTTACACCCCTGTGGTGGAGAGAGGGGCCACAGTTGAAGTG	1723
OY	1621	TGGACAGAAACAGAGATTCGCCCATTTCTGCGGGCCCTGGTCCCAACCCCAAGATTTCTCT	1680
Db	1724	TGGGCAGAGACAGAGATTCGCCCATTTGCACTGCGCTGCTGGTCCCAACCCCAAGATCTCTCT	1783
OY	1681	GCTGGATGAGGCAACGTCACGCTCTGGACACTGAAGATGAAGAGAGTGTTCAGGTGGCCCT	1740
Db	1784	GCTGGATGAGGGCCACGTCACGCTTTGGACACAGAAAGCAGAGAGTGTTCAAGGTGGCTCT	1843
OY	1741	GGATTAAGGCCAGAAAGGCCGCGACTACCATTTGATAGCTCACTGCTTTGTCTACAGTTTCG	1800
Db	1844	GGATTAAGGCCAGAAAGAGTGGGACACACATTTGATAGCTCACTGCTTTGTCTACAGTTTCG	1903
OY	1801	TAAATGCGAGTCTCTGCTGTTTGTGATGAGATCTTGGGAGAAAGGAAATCATGA	1860
Db	1904	TAAATGCTAGCTCTCTGCTGTTTGCATGTGAGATCATTTGGGAGAAAGGAAATCATGA	1963
OY	1861	TGAACCTATGAAGAGAGAGGCACTTACTTCAAATTTGTCAACATGCAAGACAAGAGAA	1920
Db	1964	TGAACCTATGAAGAGAGAGGCACTTACTTCAAACCTTGCACAAATGCAAGACAGCAGGAA	2023
OY	1921	TGAATTTGAGTTAGAAAAATGCCACTGCTGATTCCAAATATGAAATGATGCTTGGAAAT	1980
Db	2024	TGAATTTGAAATTAAGAAAAATCAGCTGATGATTCCAAATATGAAATTTGATGCTTGGAAAT	2083
OY	1981	GTCCTCAAAATATTCAGGATTCAGGTTTAAATTAAGAAACAAATCACTCGGAGGATGTACA	2040
Db	2084	GTCCTCAAAATGATTCAGAGATTCAGCTCTAATTAAGAAAAAATCAACTCGTAGGAGTGTCCG	2143
OY	2041	TGCACCAAGGCCAGACAGAAAGCTTGGTACAAAGAGGACTTGAATGAGATGTACC	2100
Db	2144	TGGAATCAAGGCCCAAGACAGAAAGCTGTAGTACCAAGAGGCTGTGGATGAAGATGTACC	2203
OY	2101	TCCAGTTTCTCTTGTGAGAGATTTCTGAAGCTGAACCTCACTGAATGGCTTATTTTGTGCT	2160

Db	2204	TCGAGTTTCCTTTGGAGGATTATGAGCTAAATTTAACTGAAAGGCTTATTTGGTGT	2263
QY	2161	TGCTATATTTTTGGCTATTATAAAGGAGGCGCTGCACACCAATTTTCAATAATTTTC	2220
Db	2264	TGGTGATTTTGGACATTATAAATGAGGCGCTGCACACCAATTTTGCATAAATATTTTC	2323
QY	2221	AAGATTATATGGGATCTTTACCCGAGATAGAGATGCTGAACCAAAAGACAGAAATCTAA	2280
Db	2324	AAAGATTATATGGGGTTTTTACAAGAAATATATATCTGTAACCAAAACAGCAAGATGTAA	2383
QY	2281	CATGTTTTCGTATATGTTTCTAGTCCCTTGGAATATTTCTTTTATATCAATTTTCTCCA	2340
Db	2384	CTTGTTTTCACTATTTGTTTCTAGCCCTTGGAAATATTTCTTTTATATACATTTTCTTCA	2443
QY	2341	GGGCTTCACATTTTGGCAAAAGCTGGGAGATTCCTCACTAAGGGCTTCGATACATGTGTTTT	2400
Db	2444	GGGTTTCACATTTTGGCAAAAGCTGGAGAGATTCCTCAACAAAGGGCTTCGATACATGTGTTTT	2503
QY	2401	CAGATTCATGCTGAGACAGGATGTCAGTCGCTGTTGATGATACCTTAATAAACACCACTGGAGC	2460
Db	2504	CCGATTCATGCTGAGACAGGATGTCAGTCGCTGTTGATGATACCTTAATAAACACCACTGGAGC	2563
QY	2461	ATTACAAACAGGGTTTCCCATATGTCGGCTCAAGTTAAAGGGCTATAGGTTCCAGGCT	2520
Db	2564	ATTACTATCCAGGCTGCGCCATATGCTGCTCAAGTTAAAGGGCTATAGGTTCCAGGCT	2623
QY	2521	TGCTGTATTAACCCAGAAATATAGCAAAATCTTGGGACAGGCAATTATATCTTAACTCTA	2580
Db	2624	TGCTGTATTAACCCAGAAATATAGCAAAATCTTGGGACAGGCAATTATATCTTCACTCTA	2683
QY	2581	TGCTTGGCATTAACACTTTTACTCTTACCAATATGACCAATCTGAATATAGAGAGAT	2640
Db	2684	TGGTTGGCACTTAACACTGTTTACTCTTACCAATATGACCAATCTGAATATAGAGAGAT	2743
QY	2641	TGTTGAATGAAATGTGTTGCTGGAACACACTGAAAGATTAAGAAAGAGCTAGAAAGAGC	2700
Db	2744	TGTTGAATGAAATGTGTTGCTGGAACACACTGAAAGATTAAGAAAGAGCTAGAAAGAGC	2803
QY	2701	TGGGAAGATTGCTACAGAAAGCCATGCAAAACTTCCGAACGTGTGTTCTTTGACTCGGGA	2760
Db	2804	TGGGAAGATTGCTACAGAAAGCCATGCAAAACTTCCGAACCGTGTGTTCTTTGACTCGAGGA	2863
QY	2761	GCAGAAATTGCAATACATGATATGACACAGAGTTTGCAGATACATACAAACCTCTTGAG	2820
Db	2864	GCAGAAATTGCAATACATGATATGATGCTCAGAGTTTGCAGATACATACAAACCTCTTGAG	2923
QY	2821	GAAAGCACACTCTTGGGGCTCATTTTCTATCACCCAGGCAATGATATTTTCTCTA	2880
Db	2924	GAAAGCACACTCTTGGGAATTACATTTTCTTACCCAGGCAATGATATTTTCTCTA	2983
QY	2881	TGCTGGCTGTTCGGGTTTGGTGGCTACTTGTGTGGCAATGAGTTCATGAACCTTTCAGGA	2940
Db	2984	TGCTGGAGTTTCCGGTTTGGAGCTACTTGTGTGGCAATGAGTTCATGAACCTTTCAGGA	3043
QY	2941	TGTTCTTTTGGTATTTCCACACTATGTGCTTGTGGCAATGAGAGTGGGGCAGGTACAGTTC	3000
Db	3044	TGTTCTTTTGAATTTTCCACTGTGTTGTTGTGGCAATGAGAGTGGGGCAGGTACAGTTC	3103
QY	3001	ATTTCCTCTCACTATGTCGCAAAAGCAAAAGATATAGCAGGCCACGTCATCATGATCATATGA	3060
Db	3104	ATTTCCTCTCACTATGTCGCAAAAGCAAAATATAGCAGGCCACATCATGATCATATTTGA	3163
QY	3061	AAAAAGCCCTCTGATTGACAGCTACAGCCCTCAGGCGCTCAAGCCAAATAGCTTGGAAAG	3120
Db	3164	AAAAAGCCCTTTGATTGACAGCTACAGCAAGCAAGGCGCTAATGCGGAAACAAATTGGAAG	3223
QY	3121	AAATGTGACATTTAATAGAGTCCGTTTCAACATATCCCATCGACACAGACATCCCGTGTCT	3180
Db	3224	AAATGTCACTTTTGGTAAGTGTGATTTCACTATCCCAACCGGACCGGACATCCCAAGTGTCT	3283
QY	3181	CCAGGGGTGAGCCTCGACAGGTGAAGAAAGGCGCAGAGCTGGGCGCTGATAGTAGAGAGTG	3240

Db 3284 TCAGGAGCTGAGCCTGGAGGTGAAGAAAGGCCAGACCGTGAGCTGTGCTGGGACAGCAGTGG 3343
QY 3241 CTGTGGAGAGACAGATGTTGTTTCAGTCTTACAGCGCTTCTATAGACCCCTTGGCTGGTTC 3300
|||||
Db 3344 CTGTGGAGAGACAGAGTGGTCCAGCTCTGTGAGCGGTTCTACAGACCCCTTGGAGGGAA 3403
QY 3301 AGTGTAAATTTATGTGGCAAAAGATTAAGACATGATGTCAGTGGCTCCAGCACACT 3360
|||||
Db 3404 AGTGTGCTTATGTGGCAAAAGAAATTAAGCGCACTGAATGTTTCAGTGGCTCCAGCACACT 3463
QY 3361 GGGCATGCTGTCTGAGAGCCCATCTGTTTGACTGCAGCATTGCCGAGAAACATGGCTA 3420
|||||
Db 3464 GGGCATGCTGTCTGAGAGCCCATCTGTTTGACTGCAGCATTGCTGAGAACATGGCTA 3523
QY 3421 TGGAGACAACAGCCGGGTCTGTATCATGAAGATTTATGACGACGACCAAGAGGCCAA 3480
|||||
Db 3524 TGGAGACAACAGCCGGGTGTGTATCATGAAGATGCTGAGGGGCACAAAGAGGCCAA 3583
QY 3481 CATACACCACTTCATCGAGACTCCCTGAGAAATACACACCCAGAGTAGAGACAAAG 3540
|||||
Db 3584 CATACATGCTTCATCGACTCATCTCCTAAATATAGCACTAAAGTAGAGACAAAG 3643
QY 3541 AACCAGCTCTGTGGTGGCCAGAAACAGCGCATGGCATAGCTCGCCCTCTGTAGACA 3600
|||||
Db 3644 AACTGAGCTCTGTGGTGGCCAGAAACAGCATTCATGCTGTGTGTAGACA 3703
QY 3601 GCCTCATATTTTGTGTTGGATGAAGCTATCATGCTCTGATATACAGAAAGTGAAGT 3660
|||||
Db 3704 GCCTCATATTTTGTGTTGGATGAAGCTATCATGCTCTGATATACAGAAAGTGAAGT 3763
QY 3661 TGTCCAGAAAGCCCTGACAAAGCCAGAGAGGCCGACCTGCAATGTGTATGCCCAAG 3720
|||||
Db 3764 TGTCCAGAAAGCCCTGACAAAGCCAGAGAGGCCGACCTGCAATGTGTATGCCCAAG 3823
QY 3721 CTGTGCACCATTCAGATGAGATTAATAGTGTGTTTCAGAAATGGCAAGTCAAGGA 3780
|||||
Db 3824 CTTGTCCACCATTCAGATGAGATTAATAGTGTGTTTCAGAAATGGCAAGTCAAGGA 3883
QY 3781 GCATGGCACATCAACAGCTGCTGGGCCAGAAAGCATCTATTTTCCATGCTAGTGT 3840
|||||
Db 3884 GCATGGCACATCAACAGCTGCTGGGCCAGAAAGCATCTATTTTCCATGCTAGTGT 3943
QY 3841 CCAGGCTGGAGCAAGCGCTAGTGAAGTGTGCCATATGAGCTGTTAAATATTTTAAAT 3900
|||||
Db 3944 CCAGGCTGGAGCAAGCGCCAGTGAAGTGTGACTGTATGAGATGTTAAATACTTTTAAAT 4003
QY 3901 ATTTGTGTAAACATGGCATTTAATCAAGTTAAAGTGAAGCTACTACTGAGAAAAC 3960
|||||
Db 4004 ATT--TGTATGATATGACATTTATTCAAAGTTAA--AGCAACACTTACAGAAAT 4056
QY 3961 ATGTAGAACTACTGTTTAACTTTCTGCTGCACTGAGATCATTTCCACCAAGTTCAG 4020
|||||
Db 4057 ATGAGAGAGTATGTGTTTAACTTTCTGCTGCACTGAGTTCAGAGTCTTC----- 4104
QY 4021 AGTCTCAGATTTTATATTAAGAACCA--AAGAACATTATCTGATGAATAAAT 4078
|||||
Db 4105 ----AGAGACTTCTGATTTAAAGAAACAGAGTGAAGACATCATCAAGTGAAGAAAT 4159
QY 4079 ACTGGTGTATATGATTAATAAATATAGATTAATCAAGTGAATTTTGTATTAAT 4138
|||||
Db 4160 CATAGTTTAACTGCATTAATAATTTATTAACAGAAATTAAGTGAATTTTAAAGATAAA 4219
QY 4139 TGTATATTTTGTGTTTATTTT-----ATTGTAATTTACTGTTGCTGAAGAT 4190
|||||
Db 4220 ATGTGTAATTTTGTGTTTATTTTCCCATTTGCACTGACTGACTGCTGCTAAGAAAT 4279
QY 4191 TATAGAAAGTGTAAAGTACTG-AAATGTTGAATAAGTGTAGCTATATTAATACTAA 4249
|||||
Db 4280 TATAGAAAGTGTAAAGTATGAATGTTTGCATTAAGTG--TCTATATTAATAACTAA 4336
QY 4250 ACTTTTATATCAA 4262
|||||
Db 4337 ACTTTCAAAAAA 4349

Search completed: December 9, 2002, 21:27:35
Job time : 1066 secs

D

QY	AGGCGCAGGTGGGANTGATTCCTGAAGAGGCCGTAAAGGAGACTGCACAGAACAATTCT	62
Dd	AGCCCGAGGTGACGATGATCCTAAGAGAGCCGTAAAGGAGGTGCACAGAAACCTTCT	115
QY	GGAATAATGGCGCAAAAAAGTAAAAAATAGAAAGAAAAGAAACCACACTGTACACA	122
Dd	GGAAAAATGGCGCAAAAAAGTAAAAA--GAGAAAGAAACAAAAAGAACCAACTGTACACA	172
QY	CGTTTGCAATGTTTCGTAATCAATGTGGCTGATAGTGTATATNTGTTGGTGGGACAA	182

Db 173 CATTGCAATGTTTGGCTATTCAAAATTTGGCTTGAATAGGTTGTAATGTTGGTGGGACAA 232
Qy 183 TGGCGCATCATCATGATGAGTGCACCTCCCTCATGATGCTGGTTTTGGAACATGA 242
Db 223 TGGCTGCATTCATCCATGAGCTGCACCTCCTCATGATGCTGGTTTTGGAACATGA 292
Qy 243 CAGATAGCTTTGCAAAATCGAGAAATTTCAAGAAACAAATTTTCCACTTAAATTAATG 302
Db 293 CAGATAGCTTTGCAAAATCGAGAAATTTCAAGAAACAAATTTTCCACTTAAATTAATG 352
Qy 303 AAAGTATTACGAACATACACAACTTTCATCAACCATCTGGAGGAGAAATGACACAGT 362
Db 353 AAAGTATTACGAACATACACAACTTTCATCAACCATCTGGAGGAGAAATGACACAGT 412
Qy 363 ATGCCATTATTACAGTGGGATCGGTGCTGCGCTGGTGGCTGCTTACATCCAGTTT 422
Db 413 ATGCCATTATTACAGTGGGATCGGTGCTGCGCTGGTGGCTGCTTACATCCAGTTT 472
Qy 423 CATTCGTGCTGCGCAGCAGAGACAGATACCTCAAAATTTGAAACAAATTTTTCATG 482
Db 473 CATTCGTGCTGCGCAGCAGAGACAGATACCTCAAAATTTGAAACAAATTTTTCATG 532
Qy 483 CATATCGGAGACAGGATTTGGCTGTTGACGTGATACGCTTGGGAGCTTAACACC 542
Db 533 CATATCGGAGACAGGATTTGGCTGTTGACGTGATACGCTTGGGAGCTTAACACC 592
Qy 543 GGCTCAGACAGATGTCTCCAAATCATGAAGAAATTTGGCAGCAAAATTTGATGTTCT 602
Db 593 GGCTCAGACAGATGTCTCCAAATCATGAAGAAATTTGGCAGCAAAATTTGATGTTCT 652
Qy 603 TTACATCATAGAACATTTTACCGGTTTTATAGTGGGTTTTACACGTGTTGGAAGC 662
Db 653 TTACATCATAGAACATTTTACCGGTTTTATAGTGGGTTTTACACGTGTTGGAAGC 712
Qy 663 TTACCCCTTGATTTGGCCATCAGCCCTGTTCTTGACTTTACCGCCCATCTGGGCAA 722
Db 713 TTACCCCTTGATTTGGCCATCAGCCCTGTTCTTGACTTTACCGCCCATCTGGGCAA 772
Qy 723 AGATACTATCTCATTTACTGATTAAGAACTCTTGGCCTATGCAAAAGCTGGAGCAGTAG 782
Db 773 AGATACTATCTCATTTACTGATTAAGAACTCTTGGCCTATGCAAAAGCTGGAGCAGTAG 832
Qy 783 CTGGAAGAGCTTACGACGAATCAGAACTGATTCCTTTGGAGGACAAAGAAAGAAC 842
Db 833 CTGGAAGAGCTTACGACGAATCAGAACTGATTCCTTTGGAGGACAAAGAAAGAAC 892
Qy 843 TTGAAGGTACAAACAAATTTAGAGAGCTTAAGGAATTTGGGATTAAGAAAGCTATCA 902
Db 893 TTGAAGGTACAAACAAATTTAGAGAGCTTAAGGAATTTGGGATTAAGAAAGCTATCA 952
Qy 903 CGGCCAATTTCTATTTGGGCGCTTCTTATTTGATCTATGATCATATGCTCTGGCTT 962
Db 953 CGGCCAATTTCTATTTGGGCGCTTCTTATTTGATCTATGATCATATGCTCTGGCTT 1012
Qy 963 TCTGTATGAGGACCTCTTGGTCTCTCCAGTGAATTTCTATTTGAGCAAGTACTACTG 1022
Db 1013 TCTGTATGAGGACCTCTTGGTCTCTCCAGTGAATTTCTATTTGAGCAAGTACTACTG 1072
Qy 1023 TCTTCTTTCTGTATTAATTTGGGCTTTTATTTGAGGAGCAATCCCAAGCAATTAAG 1082
Db 1073 TCTTCTTTCTGTATTAATTTGGGCTTTTATTTGAGGAGCAATCCCAAGCAATTAAG 1132
Qy 1083 CATTGCAAGGACGAGGAGCAGCTTATGAAATCTTCAAGATTAATTTGACAAATTAACAA 1142
Db 1133 CATTGCAAGGACGAGGAGCAGCTTATGAAATCTTCAAGATTAATTTGACAAATTAACAA 1192
Qy 1143 GCATTGACAGCTATTGGAAGATGAGACATAAACAGATTAATTTAAGGAAATTTGCAAT 1202
Db 1193 GCATTGACAGCTATTGGAAGATGAGACATAAACAGATTAATTTAAGGAAATTTGCAAT 1252
Qy 1203 TCATAAATGTTCACTTCACTTACCTTCTCGAAGAAAGATTAAGATCTTAAAGGCTCA 1262
Db 1253 TCATAAATGTTCACTTCACTTACCTTCTCGAAGAAAGATTAAGATCTTAAAGGCTCA 1312

Qy 1263 ACCTGAAGGTTTCAGAGTGGGACAGACAGTGGCGCTGTTGGGAAACAGTGGCTCGGGAGCA 1322
Db 1313 ACCTGAAGGTTTCAGAGTGGGACAGACAGTGGCGCTGTTGGGAAACAGTGGCTCGGGAGCA 1372
Qy 1323 GCACGACCGCTGCAGCTGATGACAGAGGCTCTATGACCCACAGTATGCAATGCTATTTG 1382
Db 1373 GCACGACCGCTGCAGCTGATGACAGAGGCTCTATGACCCACAGTATGCAATGCTATTTG 1432
Qy 1383 ATGGACAGGACATTAGGACCATTAATGTAAGGATCTTGGGAAATTTACTGCTGCTGA 1442
Db 1433 ATGGACAGGACATTAGGACCATTAATGTAAGGATCTTGGGAAATTTACTGCTGCTGA 1492
Qy 1443 GTCGAGGACCTGTGTTTGGCCACGACAGATTAAGTGAATTCCTATGCGCGGAAA 1502
Db 1493 GTCGAGGACCTGTGTTTGGCCACGACAGATTAAGTGAATTCCTATGCGCGGAAA 1552
Qy 1503 ATGTCACCATGATGAGATTTGACAAAGCTGTTAAGGAAGCCATAGCTATTTATCA 1562
Db 1553 ATGTCACCATGATGAGATTTGACAAAGCTGTTAAGGAAGCCATAGCTATTTATCA 1612
Qy 1563 TGAACCTAACCTAATTAATTTGACACTCTGTTGGAGAGAGGAGGCGCCAGCTGATGCTG 1622
Db 1613 TGAACCTAACCTAATTAATTTGACACTCTGTTGGAGAGAGGAGGCGCCAGCTGATGCTG 1672
Qy 1623 GACAGAAACAGAGAAATGCCATTGCTGGGCGCTGTTCCAGACCCACAGATTTCTGC 1682
Db 1673 GACAGAAACAGAGAAATGCCATTGCTGGGCGCTGTTCCAGACCCACAGATTTCTGC 1732
Qy 1683 TGGATGAGGCAAGCTCAGCTGAGCACTGGAACCTGGAAGAGAGTGGTTCAGTGGCCCTGG 1742
Db 1733 TGGATGAGGCAAGCTCAGCTGAGCACTGGAACCTGGAAGAGAGTGGTTCAGTGGCCCTGG 1792
Qy 1743 ATAAAGCCAGAAAAGGCGGAGCTACATTTGATGATGCTATGCTTCTACAGTTCTGA 1802
Db 1793 ATAAAGCCAGAAAAGGCGGAGCTACATTTGATGATGCTATGCTTCTACAGTTCTGA 1852
Qy 1803 ATGCCGATGCTATGCTGCTGTTTGTATGATGAGATCTTTGAGAAAGAAATGATG 1862
Db 1853 ATGCCGATGCTATGCTGCTGTTTGTATGATGAGATCTTTGAGAAAGAAATGATG 1912
Qy 1863 AACCTATGAAGAGAGGAGGCAATTTACTTCAAACTGTGCACATGCAAGAGAGAAATG 1922
Db 1913 AACCTATGAAGAGAGGAGGCAATTTACTTCAAACTGTGCACATGCAAGAGAGAAATG 1972
Qy 1923 AAATTGATTTAGAAAATGCCACTGCTGAATCCAAAAGTGAAGTCAAGTCAAGTATACATG 2042
Db 1973 AAATTGATTTAGAAAATGCCACTGCTGAATCCAAAAGTGAAGTCAAGTCAAGTATACATG 2092
Qy 2043 CACCAAGAGGCGAAGACAGAAAGCTTGTATCAAAAGAGAGCTTGAATGAGATGACTC 2102
Db 2093 CACCAAGAGGCGAAGACAGAAAGCTTGTATCAAAAGAGAGCTTGAATGAGATGACTC 2152
Qy 2103 CAGTTTCTTCTGAGAGATTTGAAAGCTGAACCTCAACGTAAGTGGCTTATTTGGTTG 2162
Db 2153 CAGTTTCTTCTGAGAGATTTGAAAGCTGAACCTCAACGTAAGTGGCTTATTTGGTTG 2212
Qy 2163 GTATATTTGTGTATTTAATAAGGAGGCGCTCAACACAGCATTTTCAATTAATTTTCAA 2222
Db 2213 GTATATTTGTGTATTTAATAAGGAGGCGCTCAACACAGCATTTTCAATTAATTTTCAA 2272
Qy 2223 GATTTATAGGATCTTTTACCCGAGATGAGATCTGAAAACAAACAGACAGATTAATACA 2282
Db 2273 GATTTATAGGATCTTTTACCCGAGATGAGATCTGAAAACAAACAGACAGATTAATACA 2332
Qy 2283 TGTATTTCTGTATTTCTTACAGTCTTGAATTAATTTCTTTTACATTTTCTCCACAG 2342
Db 2333 TGTATTTCTGTATTTCTTACAGTCTTGAATTAATTTCTTTTACATTTTCTCCACAG 2392

QY 2343 GCTTCACATTTGGCAAGCTGGGAGATTCCTACATAGCGGCTTCGATACATGGTTTCA 2402
|||||
Db 2393 GCTTCACATTTGGCAAGCTGGGAGATTCCTACATAGCGGCTTCGATACATGGTTTCA 2452
QY 2403 GATCCATGCTGAGACAGGATGTCAGCTGGTTTGATGACCTTAAAAACACACTGAGCAT 2462
|||||
Db 2453 GATCCATGCTGAGACAGGATGTCAGCTGGTTTGATGACCTTAAAAACACACTGAGCAT 2512
QY 2463 TGAACACACAGGCTTGGCCAAATGATGGGCTCAAGTTAAAGGGGCTTATAGTCCAGGCTG 2522
|||||
Db 2513 TGAACACACAGGCTTGGCCAAATGATGGGCTCAAGTTAAAGGGGCTTATAGTCCAGGCTG 2572
QY 2523 CTGTCAATTACCCAGAAATATAGCAAAATCTTGGACAGGCAATTATATCTTAAATCATG 2582
|||||
Db 2573 CTGTCAATTACCCAGAAATATAGCAAAATCTTGGACAGGCAATTATATCTTAAATCATG 2632
QY 2583 GTTGGCAATTACATCTTTTACTCTTACCAATTTGACCAATCTATGTCATAGCAGAGATG 2642
|||||
Db 2633 GTTGGCAATTACATCTTTTACTCTTACCAATTTGACCAATCTATGTCATAGCAGAGATG 2692
QY 2643 TTGAAATGAAATGTTGCTCTGAGACAGACTGAATAATGAAGAGCTAAGAGAGCTG 2702
|||||
Db 2693 TTGAAATGAAATGTTGCTCTGAGACAGACTGAATAATGAAGAGCTAAGAGAGCTG 2752
QY 2703 GGAAGATTGCTACAGAAAGCCATCGAAACCTTGGAACTGTTGTTCTTGAAGCTGGAGC 2762
|||||
Db 2753 GGAAGATTGCTACAGAAAGCCATCGAAACCTTGGAACTGTTGTTCTTGAAGCTGGAGC 2812
QY 2763 AGAAGTTTAAATGATGATGACAGAGTTTGCATATACATACAGAAACCTTGGAGC 2822
|||||
Db 2813 AGAAGTTTAAATGATGATGACAGAGTTTGCATATACATACAGAAACCTTGGAGC 2872
QY 2823 AAGCAACACATCTTGGGCTCATATTTCTATGACCCAGGCAATGATATTTTCCATG 2882
|||||
Db 2873 AAGCAACACATCTTGGGCTCATATTTCTATGACCCAGGCAATGATATTTTCCATG 2932
QY 2883 CTGGCTGTTCCGATTTGGTGGCTTACTTGGTGGCAATGATGATCAATCTTCCAGATG 2942
|||||
Db 2933 CTGGCTGTTCCGATTTGGTGGCTTACTTGGTGGCAATGATGATCAATCTTCCAGATG 2992
QY 2943 TTCTTTGGTATTTCCAGCTATTTGCTTTGGTGGCAATGATGATCAATCTTCCAGATG 3002
|||||
Db 2993 TTCTTTGGTATTTCCAGCTATTTGCTTTGGTGGCAATGATGATCAATCTTCCAGATG 3052
QY 3003 TTGCTCTGATCTATGCTTAAAGCCAAATATGACAGCCAGCTCATGATCATTTTAA 3062
|||||
Db 3053 TTGCTCTGATCTATGCTTAAAGCCAAATATGACAGCCAGCTCATGATCATTTTAA 3112
QY 3063 AAAAGCCCTGATGATGACAGCTACAGCCCTCAGCGGCTCAAGCCAAATACGTTGGAAGAA 3122
|||||
Db 3113 AAAAGCCCTGATGATGACAGCTACAGCCCTCAGCGGCTCAAGCCAAATACGTTGGAAGAA 3172
QY 3123 ATGTGACATTTAAAGAGGCTGTTCAATATCCCACTGACAGACATGCTCCGCTGCTC 3182
|||||
Db 3173 ATGTGACATTTAAAGAGGCTGTTCAATATCCCACTGACAGACATGCTCCGCTGCTC 3232
QY 3183 AGGGGCTGAGCTCGAGCTGAAGAAGGCGCAGAGCTGGCCCTCGTAGATGACAGTGGCT 3242
|||||
Db 3233 AGGGGCTGAGCTCGAGCTGAAGAAGGCGCAGAGCTGGCCCTCGTAGATGAGTGGCT 3292
QY 3243 GTGGGAAGAGACAGATTTGACGCTCCTAGAGGCTTCTATGACCCCTTGGCTGGTTCAG 3302
|||||
Db 3293 GTGGGAAGAGACAGATTTGACGCTCCTAGAGGCTTCTATGACCCCTTGGCTGGTTCAG 3352
QY 3303 TTGCTATGATGAGCAAAAGATTAAGCACTGAATTCAGAGGCTCCGAGACACAGCTGG 3362
|||||
Db 3353 TTGCTATGATGAGCAAAAGATTAAGCACTGAATTCAGAGGCTCCGAGACACAGCTGG 3412
QY 3363 GCATGCTGCTCAGAGCCCATCTGTTGATGACATGACATGCTCCGAGAACTTCTATG 3422
|||||
Db 3413 GCATGCTGCTCAGAGCCCATCTGTTGATGACATGACATGCTCCGAGAACTTCTATG 3472
QY 3423 GAGACAAAGCCGGGCTGCTATCAGATGAAGATTTATGACAGGACAGCAAGAGGCGCAAC 3482

Db 3473 GAGACAAAGCCGGGCTGCTATCAGATGAAGATTTATGAGGCGCCAAAGAGCCAAAC 3532
|||||
QY 3483 TACACACCTTCATGAGACACACTCCCTGAGAAATATACACACAGATGAGAGCAAGAGAA 3542
|||||
Db 3533 TACACACCTTCATGAGACACACTCCCTGAGAAATATACACACAGATGAGAGCAAGAGAA 3592
QY 3543 CCCAGCTCTGAGGCGAGAAAGAGGCGCATTTGATAGCTGCGGCTGTTAGACAGC 3602
|||||
Db 3593 CCCAGCTCTGAGGCGAGAAAGAGGCGCATTTGATAGCTGCGGCTGTTAGACAGC 3652
QY 3603 CTCATATTTTGGCTTTTGGATGAAAGCTTACATGACTCTGGATTCAGAAAGTGAAGGTTG 3662
|||||
Db 3653 CTCATATTTTGGCTTTTGGATGAAAGCTTACATGACTCTGGATTCAGAAAGTGAAGGTTG 3712
QY 3663 TCCAAAGAGCCCTGAGCAAAAGCCAGAAAGGCGGCACTGGATTTGATGCGCCACGCT 3722
|||||
Db 3713 TCCAAAGAGCCCTGAGCAAAAGCCAGAAAGGCGGCACTGGATTTGATGCGCCACGCT 3772
QY 3723 TGTCCACCATCCAGAAATGACATTTATATGTTGTTTCAAAATGCGCAAGTCAAGGAGC 3782
|||||
Db 3773 TGTCCACCATCCAGAAATGACATTTATATGTTGTTTCAAAATGCGCAAGTCAAGGAGC 3832
QY 3783 ATGGCACACATCAACAGCTGCTGCGCCAGAAAGGCACTATTTTTCATGCTGAGTCC 3842
|||||
Db 3833 ATGGCACACATCAACAGCTGCTGCGCCAGAAAGGCACTATTTTTCATGATGAGTCC 3892
QY 3843 AGGCTGGAGCAAGGCGCTAGTGAACCTGGGCAATATGAGCTGTTTAAATATTTTAAATAT 3902
|||||
Db 3893 AGGCTGGAGCAAGGCGCTAGTGAACCTGGGCAATATGAGCTGTTTAAATATTTTAAATAT 3952
QY 3903 TTGTGTTAAACATGAGCAATTAATCAAGTTAAAGGAGCACTTCTGGAAGAACTAT 3962
|||||
Db 3953 TTGTGTTAAACATGAGCAATTAATCAAGTTAAAGGAGCACTTCTGGAAGAACTAT 4012
QY 3963 GTAGAACTACCTGTTTAACTTTCTGCTGCAACTGAAATCATCTTCCACAGTTCAGAG 4022
|||||
Db 4013 GTAGAACTACCTGTTTAACTTTCTGCTGCAACTGAAATCATCTTCCACAGTTCAGAG 4072
QY 4023 TCTTCAGATTTTAAATTAATTAAGAACCAAAAGAACATTTCTGATGAATTAATTAAT 4082
|||||
Db 4073 TCTTCAGATTTTAAATTAATTAAGAACCAAAAGAACATTTCTGATGAATTAATTAAT 4132
QY 4083 GTGTATATTCATTTAAATTAATTAAGATTAATCAAGATTAATTTGTTATTAATTTGTA 4142
|||||
Db 4133 GTGTATATTCATTTAAATTAATTAAGATTAATCAAGATTAATTTGTTATTAATTTGTA 4192
QY 4143 TAATTTTGTATTAATTTTAACTTACTGCTTGTGCGAAGATTAATTAAGATTTGTA 4202
|||||
Db 4193 TAATTTTGTATTAATTTTAACTTACTGCTTGTGCGAAGATTAATTAAGATTTGTA 4252
QY 4203 AAAAGTACTGAATGTTTGAATTAAGGCTAGCTATTAATTAAGCACTTTTATATCAA 4262
|||||
Db 4253 AAAAGTACTGAATGTTTGAATTAAGGCTAGCTATTAATTAAGCACTTTTATATCAA 4312
QY 4263 AAAAA 4267
|||||
Db 4313 AAAAA 4317

RESULT 2
US-10-072-621-2
; Sequence 2, Application US/10072621
; Patent No. US20020169137A1
; GENERAL INFORMATION:
; APPLICANT: Reiner, Peter B.
; APPLICANT: Connop, Bruce P.
; APPLICANT: Pollard, Michelle
; TITLE OF INVENTION: REGULATION OF AMYLOID PRECURSOR PROTEIN EXPRESSION
; FILE REFERENCE: 100103.402
; CURRENT APPLICATION NUMBER: US/10/072,621
; CURRENT FILING DATE: 2002-02-08

Db 2380 GTCCTCAATGATTCAGATCCAGTCTAATAGAAAAAGATCACTCGTAGAGTGTCCG 2439
Qy 2041 TGCACCAAGGCCAAGACAGAAAGCTTGGTACAAAAAGAGACTGATGAGTATGACC 2100
Db 2440 TGGATCACAAGGCCAAGACAGAAAGCTTAGTACCAAAAGAGGCTGTGATGAAGATATACC 2499
Qy 2101 TCCAGTTTCCCTCTGAGAGATTTCTGAGCTGACAGTCAAGTAAAGGCTTATTTTGTGTGT 2160
Db 2500 TCCAGTTTCCCTTTGAGAGATTAAGAGCTTAATTTAGCTGAATTTGAGCTTATTTGTTGT 2559
Qy 2161 TGGTATATTTTGTCTATTAATAAAGAGCCCTGCACACAGCAATTTCAATATATTTTC 2220
Db 2560 TGGTATATTTTGTCTATTAATAAAGAGCCCTGCACACAGCAATTTCAATATATTTTC 2619
Qy 2221 AAGGATTTATAGGATCTTTACCCGATGAGATCTTGAAGAACAAAGCAGCAATATGTA 2280
Db 2620 AAGGATTTATAGGGGTTTTACAGAAATTTGATGATCTGAAACAAAGCAGCAATATGTA 2679
Qy 2281 CATGTTTCCGATATGTTTCTAGTCTTGGATTAATTTCTTTATTAATTTTCTTCA 2340
Db 2680 CTGTTTCTACTATTTGTTCTAGCCCTTGGATTAATTTCTTTATTAATTTTCTTCA 2739
Qy 2341 GGGCTTACATTTGGCAAAAGCTGGAGATCTCTACATTAAGCGCTTCATACATGTTTT 2400
Db 2740 GGGTTTACATTTGGCAAAAGCTGGAGATCTCTACACCAAGCGCTTCATACATGTTTT 2799
Qy 2401 CAGATCCATGCTGAGACAGAGATGTCAGCTGTTGATGACCCCTAAACACACACTGGAGC 2460
Db 2800 CCGATCCATGCTGAGACAGAGATGATGATGTTGATGACCCCTAAACACACACTGGAGC 2859
Qy 2461 ATTGACACAGGCTTGGCAATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2520
Db 2860 ATTGACTACCAAGCTTGGCAATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2919
Qy 2521 TGTCTGATATACCCAGATATAGCAAAATCTTGGACAGCATTATATATCTTAACTTA 2580
Db 2920 TGTCTGATATACCCAGATATAGCAAAATCTTGGACAGCATTATATATCTTAACTTA 2979
Qy 2581 TGGTGGCAATTAACACTTTACTCTTACCAATTTGATGATGATGATGATGATGATGATGAT 2640
Db 2980 TGGTGGCAATTAACACTTTACTCTTACCAATTTGATGATGATGATGATGATGATGATGAT 3039
Qy 2641 TGGTGGCAATTAACACTTTACTCTTACCAATTTGATGATGATGATGATGATGATGATGAT 2700
Db 3040 TGGTGGCAATTAACACTTTACTCTTACCAATTTGATGATGATGATGATGATGATGATGAT 3099
Qy 2701 TGGGAAGATTTGCTACAGAGACCATCGAAACCTTCCAACTGTTTCTTCTGACTCGGGA 2760
Db 3100 TGGGAAGATTTGCTACAGAGACCATCGAAACCTTCCAACTGTTTCTTCTGACTCGGGA 3159
Qy 2761 GCAGAGTTTGAATACATGATATGCACAGAGTTTGCAGAGTACATACAGAAACTCTTTGAG 2820
Db 3160 GCAGAGTTTGAATACATGATATGCACAGAGTTTGCAGAGTACATACAGAAACTCTTTGAG 3219
Qy 2821 GAAGAGACATCTTGGGGTCTCATTTTCTATCACCCGAGGAAATGATGATGATGATGATGAT 2880
Db 3220 GAAGAGACATCTTGGGGTCTCATTTTCTATCACCCGAGGAAATGATGATGATGATGATGAT 3279
Qy 2881 TGTGCGCTGTTTCCGTTTGTGTGCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2940
Db 3280 TGTGCGCTGTTTCCGTTTGTGTGCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 3339
Qy 2941 TGTCTTTTGTGATTTCTGAGCTATTTGTTTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 3000
Db 3340 TGTCTTTTGTGATTTTCTGAGCTGTTTGTGCTTGTGCTGCTGCTGCTGCTGCTGCTGCTGCT 3399
Qy 3001 ATTGCTTCCGATTTGCAAGGCCAAGATATGAGAGGCCGAGCTATCATGATCATGATGTA 3060
Db 3400 ATTGCTTCCGATTTGCAAGGCCAAGATATGAGAGGCCGAGCTATCATGATCATGATGTA 3459
Qy 3061 AAAAAGCCCTCTGATTTGACAGCTACAGCCCTCAGCGCTCAAGCCAAATACGTTGGAAG 3120
Db 3460 AAAAAGCCCTTTGATTTGACAGCTACAGCCGCTCAAGGCTCAATGCTCAAGCTCAAGGAG 3519

Qy 3121 AATGTGACATTTAATGAGTGTGTTCAACTATCCAGCTGACAGACATCCCGTGTCT 3180
Db 3520 AATGTGACATTTAATGAGTGTGTTCAACTATCCAGCTGACAGACATCCCGTGTCT 3579
Qy 3181 CCAGGAGCTGAGCTTCCAGAGTGAAGAGGCGGAGGCGGCTGAGAGTATGAGTATG 3240
Db 3580 TCGAGGAGCTGAGCTTCCAGAGTGAAGAGGCGGAGGCGGCTGAGAGTATGAGTATG 3639
Qy 3241 CTGTGGAAAGACACAGTTGTGTCTGAGGCGCTTATGAGCCCTTGTGCTGCTGCT 3300
Db 3640 CTGTGGAAAGACACAGTTGTGTCTGAGGCGCTTATGAGCCCTTGTGCTGCTGCTGCT 3699
Qy 3301 AGTGTAAATTTGAGCAAGAGATTAAGACAGCTGATGCTAGTGGCTCCGAGCAGCT 3360
Db 3700 AGTGTGCTGCTGAGCAAGAGATTAAGACAGCTGATGCTGAGTGGCTCCGAGCAGCT 3759
Qy 3361 GGGCATGCTGCTGAGAGAGCCCATCTGTTTACCTGAGCAGTATGCTGAGAGCAATGCTTA 3420
Db 3760 GGGCATGCTGCTGAGAGAGCCCATCTGTTTACCTGAGCAGTATGCTGAGAGCAATGCTTA 3819
Qy 3421 TGGAGCAAAAGCCGCGGTCTGATACATGAAGATTTGACAGCAGCAGCAGAGCCAA 3480
Db 3820 TGGAGCAAAAGCCGCGGTCTGATACAGAGAGATTTGAGAGGAGCAGAGAGCCAA 3879
Qy 3481 CATACACACTGATGAGACACTCCCTGAGAAATACAAACACAGATGAGAGCAAAAG 3540
Db 3880 CATACACTGCTTCACTGAGTACTGCTTAATTAATATACACTTAAGATGAGAGCAAAAG 3939
Qy 3541 AACCGAGCTCTGCTGAGCCAGAAAGCCCATTTGCTAGCTGCTGCTGCTGCTGCTGCT 3600
Db 3940 AACCGAGCTCTGCTGAGCCAGAAAGCCCATTTGCTAGCTGCTGCTGCTGCTGCTGCT 3999
Qy 3601 GCTCTATTTTGTCTTTTGGATGAGCTACATGCTGCTGCTGCTGCTGCTGCTGCTGCT 3660
Db 4000 GCTCTATTTTGTCTTTTGGATGAGCTACATGCTGCTGCTGCTGCTGCTGCTGCTGCT 4059
Qy 3661 TGTCCAAAGAGCCCTGAGCAAAAGCCAGAGAGGCGCACCTGCTGATGCTGCCACCG 3720
Db 4060 TGTCCAAAGAGCCCTGAGCAAAAGCCAGAGAGGCGCACCTGCTGATGCTGCCACCG 4119
Qy 3721 CTGTGCAACATCTGAGATGAGTAAATGATGATGATGATGATGATGATGATGATGATGAT 3780
Db 4120 CTGTGCAACATCTGAGATGAGTAAATGATGATGATGATGATGATGATGATGATGATGAT 4179
Qy 3781 GCATGGCACATTCACAGCTGCTGCTGCCAGAAAGCATCTATTTTCTATGCTCACTGT 3840
Db 4180 GCATGGCACATTCACAGCTGCTGCTGCCAGAAAGCATCTATTTTCTATGCTCACTGT 4239
Qy 3841 CCAGGCTGAGCAAGCCGCTAGTGAACCTGAGCAATGAGCTGTTAAATATTTTAAAT 3900
Db 4240 CCAGGCTGAGCAAGCCGCTAGTGAACCTGAGCAATGAGTATGATGATGATGATGATGAT 4299
Qy 3901 ATTGTTTAAACATGCAATTAATTAACAGTAAAGAGTGGCCTACCTAGGAAACCT 3960
Db 4300 ATT -TGTTTGAATATGATTAATTAATTAACAGTAAAGAGTGGCCTACCTAGGAAAT 4352
Qy 3961 AGTGTAGACTGCTGTTTACATTTCTGCTGCACTGGAAGTATTTCCACCAAGTTGAG 4020
Db 4353 AGTGTAGACTGCTGTTTACATTTCTGCTGCACTGGAAGTATTTCCACCAAGTTGAG 4400
Qy 4021 AGTGTAGACTGTTTAAATTAAGAAACAA -AAGAAACATTTCTGATGAGTAAAT 4078
Db 4401 -----AGAGACTGCTATTAATTAAGAAACAGAGTGAAGACATCAAGTGAAGAAAT 4455
Qy 4079 ACTGTTTAAATGCAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 4138
Db 4456 CATAGTTTAAATGCAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 4515
Qy 4139 TGTATATATTTTGTATATTTT -----ATTGTACTTACTGCTTGTGCAAGAT 4190
Db 4516 ATGTGTATTTTGTATATTTTCCATTTTGAACGTETACGTGCTGCTTAAGAAAT 4575

QY	2346	TCACATTTGGCAAAAGCTGGGGAGATTCCTCACTAAGCGGGCTTGATACATGCTTTTCACAT	2
Db	3543	TCACATTTGGCAAAAGCTGGAGATTCCTCAACAAAGCGGCTCCGATACATGATGTTTCCGAT	3
QY	2406	CCATCTCAGACAGATGTCAGCTGGCTTTATGATACCTTAATAAACACCACTGGAGATTA	2
Db	3603	CCATGCTCAGACAGATGTCAGTGGTTTATGATACCTTAATAAACCACTGGAGATTA	3
QY	2466	CAACGAGGCTTGCCAAATGATGGGCTCAAGTAAAGGGCTATAGTTCAGAGCTTGC	2
Db	3663	CTACGAGGCTTGCCAAATGATGGTGGTTTATGATACCTTAATAAACCACTGGAGATTA	3
QY	2526	TCATTTACCCAGAAATATGCAAACTTGGGACAGCATTTATATCTTAACTATGCTT	2
Db	3723	TAATTTACCCAGAAATATGCAAACTTGGGACAGCATTTATATCTTAACTATGCTT	3
QY	2586	GGCAATTTAACATTTTACTCTTGTAGCAATTTGTCCCATATGSCAATPACAGAGATTGTTG	2
Db	3783	GGCAATTTAACATCTGTACTCTTATAGCAATTTGTCCCATATGSCAATPACAGAGATTGTTG	3
QY	2666	AAATGAAATAGTTGTGTGACACAAGCACTGAAAGATMAAAGAGCTGAAGAGACTGGGA	2
Db	3843	AAATGAAATAGTTGTGTGACACAAGCACTGAAAGATMAAAGAGCTGAAGAGACTGGGA	3
QY	2706	AGATTGCTCAGAAAGCAATCGAAACTTCCGAACTGTTGTTCTTTACTGCGGAGACA	2
Db	3903	AGATTGCTCAGAAAGCAATCGAAACTTCCGAACTGTTGTTCTTTACTGCGGAGACA	3
QY	2766	AGTTTGATATACATGTATGCACAGAGTTTGCAGATACATACAGAACTCTTTAGGAAG	2
Db	3963	AGTTTGATATATGTATGTCTGAGATTTGCAGATACATACAGAACTCTTTAGGAAG	40
QY	2826	CACACATCTCGGGGCTCAATTTCTATACCCAGGCAATGATGTATTTTCTATGCTG	28
Db	4023	CACACATCTTTGGAATTCATTTTCTTCAACCAAGGCAATGATGTATTTTCTATGCTG	40
QY	2886	GCTGTTTCCGTTTGGTGCCCTACTTGGTGGCAATGATGTATGACTTTCAGAGATGTC	29
Db	4083	GATGTTTCCGTTTGGAGCCACTTGGTGGCACATTAACATCAGACTTTCAGAGATGTC	41
QY	2946	TTTTGGTATTTTCAGTATTTGTCTTTGGTCCCATGGCAATGGGCGAGCTCAATCATTTG	30
Db	4143	TGTTAGTATTTTCAAGCTGTGTGCTTTTGGTCCCATGGGCGAGCTCAATCATTTG	42
QY	3006	CTCTGACTATGCCAAAGCCAAAGTATCAGACAGCCACTCATCATGATCAATTGAAAAA	30
Db	4203	CTCTGACTATGCCAAAGCCAAATATCAGACAGCCCATATCATGATCAATTGAAAAA	42
QY	3066	GCCCTCGATTGACGTCACAGGCCCTCAGGCGCTCAAGCCAAATATCGTTGAAAGAAATG	31
Db	4263	CCCTTTGATTGACACTACAGCCAGCGGAAGGCTATGCCCAACATTTGAAAGAAATG	43
QY	3126	TGACATTTAATGAGTGTGTTCAACTATCCCACTCGACAGACATCCCGTGGCTCAGG	31
Db	4323	TCACATTTGATGTAAGTTGTATTCACATATCCCAACCGAGCATCCCAAGTGGTCAAG	43
QY	3186	GGGTGAGCCCTCGAGGGGAAGAGGGCCAGACAGCTGGCCCTGTAGTAGAGATGGCTGTG	32
Db	4383	GACTGAGCCCTCGAGGTAGAGAGAGGGCCAGACAGCTGGGCTGTAGTGGCAGAGTGGCTG	44
QY	3246	GGAAAGACACAGTTTTCAGCTCTAGAGCCCTTCTATGACCCCTTGGGCTGATCTAGTC	33
Db	4443	GGAAAGACACATGTGTCCAGCTCTGAGCGGTTCTAGACCCCTTGGCAGGAAAAATGCT	45
QY	3306	TAATTGATGGCAAAAGATTAAGACCCGAAATGTCCAGTGGCTTCGAGACACACTGGGCA	33
Db	4503	TGCTTGATGGCAAAATTAAGGACTGTATGTAGTGGCTTCGAGACACACTGGGCA	45
QY	3366	TGCTGTCTCAGAGCCCATCTGTTTACTGACAGCATGGCCGAGAACATTTGGCTTATGAG	34
Db	4563	TGCTGTCCCAAGAGCCCATCTGTTTACTGACAGCATTTGCTAGAACATTTGGCTTATGAG	46
QY	3426	ACAACAGCGCGGTGTATCAACATGAAGACATTTATGAGCGACCAAGAGGCAACATATAC	34

```
Db 4623 ACACAGCCGGGTGTGTACAGAGAAAGATCGTAGGCGCAGCAAGAGCCCAACATAC 4682
OY 3486 ACCACTTATGAGACACCTCCCTGAGAAATACACACAGATGAGACAAAGAACCC 3545
Db 4683 ATGCTTCATGAGTACACTGCTTAATATATAGCACTAAAGTAGAGACAAAGAACTC 4742
OY 3546 ACCTCTGAGTGGCCAGAAACAGCATTTGCGATAGCTGCGGCTCTTTAGACAGCTC 3605
Db 4743 ACCTCTGAGTGGCCAGAAACAGCATTTGCGATAGCTGCGGCTCTTTAGACAGCTC 4802
OY 3606 AATTTTCTTTTGTGATGAAAGCTACATCGCTCTGATACAGAAAGTGAAGCTTCC 3665
Db 4803 AATTTTCTTTTGTGATGAAAGCTACATCGCTCTGATACAGAAAGTGAAGCTTCC 4862
OY 3666 AAGAAAGCCCTGACAAAGCCAGAGAGAGCCGACCTGCTATGTATGATGCGCCAGCTTGT 3725
Db 4863 AAGAAAGCCCTGACAAAGCCAGAGAGAGCCGACCTGCTATGTATGATGCGCCAGCTTGT 4922
OY 3726 CCACCATCCAGATGAGATTTAATAGTGTGTTTCAAGATGCGAAAGTCAAGAGCATG 3785
Db 4923 CCACCATCCAGATGAGATTTAATAGTGTGTTTCAAGATGCGAAAGTCAAGAGCATG 4982
OY 3786 GCACACATGACACAGCTGCTGCGCCAGAAAGCATTTATTTTCCATGCTAGTCCAGG 3845
Db 4983 GCACACATGACACAGCTGCTGCGCCAGAAAGCATTTATTTTTCATGCTAGTCCAGG 5042
OY 3846 CTGAGCAAAAGCGCTAGTGAAGTGTGCGCCATGAGAGCTGTTAAATATTTTAAAT 3900
Db 5043 CTGAGCAAAAGCGCGAGTGAAGTGTGCTAGTATGATGATTTAAATCTTTTAAAT 5097

RESULT 4
US-09-306-417-1
Sequence 1, Application US/09306417
Patent No. US20020103144A1
GENERAL INFORMATION:
APPLICANT: Heinrich-Pette-Institut
TITLE OF INVENTION: Retroviral Gene Transfer Vectors
FILE REFERENCE: P50491
CURRENT APPLICATION NUMBER: US/09/306,417
CURRANT FILING DATE: 1999-05-06
EARLIER APPLICATION NUMBER: DE 198 22 115
EARLIER FILING DATE: 1998-05-08
NUMBER OF SEQ. ID NOS.: 17
SOFTWARE: PatentIn Ver. 2.0
SEQ. ID NO. 1
LENGTH: 8630
TYPE: DNA
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: Description of Artificial Sequence: proviral
FEATURE:
NAME/KEY: misc_feature
LOCATION: (161)..(677)
OTHER INFORMATION: 5'-LTR
FEATURE:
NAME/KEY: 5'UTR
LOCATION: (532)..(1219)
FEATURE:
NAME/KEY: mat_peptide
LOCATION: (1320)..(5062)
OTHER INFORMATION: m4 mdr-1 cDNA
FEATURE:
NAME/KEY: misc_feature
LOCATION: (5215)..(5774)
OTHER INFORMATION: 3'-LTR
FEATURE:
```

```
; NAME/KEY: misc_feature
; LOCATION: (5775)..(8630)
; FEATURE:
; OTHER INFORMATION: Plasmid backbone (pUC)
; NAME/KEY: misc_feature
; LOCATION: (1)..(8630)
; OTHER INFORMATION: retroviral expression vector Sfbeta71m4
US-09-306-417-1
```

```
Query Match 75.3%; Score 3223.4; DB 10; Length 8630;
Best Local Similarity 89.4%; Pred. No. 0;
Matches 3483; Conservative 0; Mismatches 406; Indels 6; Gaps 1;
```

```
OY 6 GCGAGTCCGGATGATTCCTGAAAGAGGCGCTAAGGGAGTGCAGAGAGAACTTCTGCA 65
Db 1209 GCGGCGCGCCATGATGATCTTGAAGGGGACCCCATGAGAGAGCAAGAAAGAACTTTT 1268
OY 66 AAGTGGGCAAAAAGTAAAAAATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 125
Db 1269 TTAACCTGAAACATTAAGAGTGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1328
OY 126 TTAGAATGTTTCCGTATTCATTAATGGCTTGAATGATGATGATGATGATGATGATG 185
Db 1329 TTTAAATGTTTCCGTATTCATTAATGGCTTGAATGATGATGATGATGATGATGATG 1388
OY 186 CTGCATCATCATGAGAGCTGCACCTCCCTCATGATGCTGATGATGATGATGATGATGATG 245
Db 1389 CTGCATCATCATGAGAGCTGCACCTCCCTCATGATGCTGATGATGATGATGATGATG 1448
OY 246 ATAGCTTGGCAATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 305
Db 1449 ATATCTTTGCAATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1502
OY 306 GTATTAGCAACATTAACCAACATTTTATCATCAACCATCTGAGAGAGAGAGAGAGAG 365
Db 1503 GAAGTGAATCAATGATPACAGAGGTTCTTCATGAAATGAGAGAGAGAGAGAGAGAG 1562
OY 366 CCTATTATACAGTGGGATCGGTGCTGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 425
Db 1563 CCTATTATACAGTGGGATCGGTGCTGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1622
OY 426 TCTGTGCTCGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 485
Db 1623 TTTGGTCCCTGGAGCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1682
OY 486 TCATGCAACAGAGAGAGTGGCTGTTGACGTGACGTGACGTGAGAGAGAGAGAGAGAG 545
Db 1683 TAAATGCAAGAGAGAGATGAGCTGTTGATGTCACGATGTTGGGAGCTTAACACCCG 1742
OY 546 TCACAGAGATGCTCCCAAAATCAATGAAGAGATGAGCAAAATTTGGAATGTTCTTTC 605
Db 1743 TTAAGATGATGCTCTTAAGATTAATGAAGTATTTAGTGTGACAAATTTGGAATGTTCTTTC 1802
OY 606 ACTCAATAGAGACATTTTTCACCGGTTTTATAGGCGTTTACAGTGTGAGAGAGAGAG 665
Db 1803 AGTCAATGCAACATTTTTCACAGGTTTATAGAGATTTACAGTGTGAGAGAGAGAG 1862
OY 666 CCCTTGTGATTTTGGCCATCAGCCCTGTTTGGCATTTTAAAGCCCATCTGCGCAAGA 725
Db 1863 CCCTTGTGATTTTGGCCATCAGCTGCTGTTTGGCATGAGTGAAGTGTGCTGCGCAAGA 1922
OY 726 TACTATCTTATTTAATGATTAAGAACTTGGCTTATGCAAAAGCTGAGAGAGAGAGAG 785
Db 1923 TACTATCTTATTTAATGATTAAGAACTTATGAGAACTTATGAGAGAGAGAGAGAGAG 1982
OY 786 AAGAACTCTAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 845
Db 1983 AAGAGGCTCTGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2042
OY 846 AAGAGTCAACAAAATTTAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 905
Db 2043 AAGAGTCAACAAAATTTAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2102
```


Db 4263 CCCCTTTGATGACAGTACAGCAGGAGGCTTAATGCCAGACATTTGGAAGAATG 4322
Qy 3126 TGCACATTAATAGAGTCGCTGTTCACTATCCACTCGACACAGACATCCCGTGTCCAGG 3185
Db 4323 TCACATTTGTAATGTTGATTAACATATCCACCCGACCGACATCCAGTGTCCAGG 4382
Qy 3186 GGGTGAAGCTCGAGTGAAGAAGGCGCAGAGCTGGCCCTGTAGTGAAGTGAAGTGTG 3245
Db 4383 GACTGAGCTCGAGTGAAGAAGGCGCAGAGCTGGCTGTGGTGGCAGCAGTGTGTG 4442
Qy 3246 GGAAGAGCAGAGTGTGTCAGTCTAGAGCCGCTTATGACCCCTTGGCTGTGACGTC 3305
Db 4443 GGAAGAGCAGAGTGTGTCAGTCTAGAGCCGCTTATGACCCCTTGGCAGGAAGTGC 4502
Qy 3306 TAATTTGATGGAAGAAGATTAAGACCTGTAATGTCCAGTGGCTCGGACACACCTGGGCA 3365
Db 4503 TGGTGTATGGAAGAAGATTAAGACCTGTAATGTCCAGTGGCTCGGACACACCTGGGCA 4562
Qy 3366 TCGTGTCTCAGAGACCCATCCCTGTTGACGAGCATTTGCTGAGAACATTTGCTATGAG 3425
Db 4563 TCGTGTCTCAGAGACCCATCCCTGTTGACGAGCATTTGCTGAGAACATTTGCTATGAG 4622
Qy 3426 ACAACAGCCGGGTCGTATCATGAAAGATTAATGCAAGCAGCAGCAGGAGGCAACATAC 3485
Db 4623 ACAACAGCCGGGTCGTATCATGAAAGATTAATGCAAGCAGCAGGAGGCAACATAC 4682
Qy 3486 ACCACTTCATGAGACACTCCCTGGAAGATTAATGCAAGCAGCAGGAGGCAACATAC 3545
Db 4683 ATGCCCTCATGAGTCACTGCTTAATTAATTAATGCACTAATTAATGAGGAGCAAGAACTC 4742
Qy 3546 AGCTCTCTGTCGAGCAAGAACAGGCAATTCAGTCCGCTCTTTGTTAGCAGCCTC 3605
Db 4743 AGCTCTCTGTCGAGCAAGAACAGGCAATTCAGTCCGCTCTTTGTTAGCAGCCTC 4802
Qy 3606 ATATTTCTCTTTGATGATGACATCACTGATGCTGATGATGATGATGATGATGATGAT 3665
Db 4803 ATATTTCTCTTTGATGATGACATCACTGATGCTGATGATGATGATGATGATGATGAT 4862
Qy 3666 AAGAAAGCCCTGGAAGAACAGGCAAGGCGCAGCTGATGATGATGATGATGATGATGAT 3725
Db 4863 AAGAAAGCCCTGGAAGAACAGGCAAGGCGCAGCTGATGATGATGATGATGATGATGAT 4922
Qy 3726 CCACCATCCAGATGAGATTTAATAGTGTGTTTCAAGATGGAAGTGAAGAGCATG 3785
Db 4923 CCACCATCCAGATGAGATTTAATAGTGTGTTTCAAGATGGAAGTGAAGAGCATG 4982
Qy 3786 GCACGATCAAGAGTGTGTCGAGGCGCAGAAAGGATCTATTTTTCATGCTCAGG 3845
Db 4983 GCACGATCAAGAGTGTGTCGAGGCGCAGAAAGGATCTATTTTTCATGCTCAGG 5042
Qy 3846 CTGAGCAGAAAGCCTTATGATGCTGATGCTGATGCTGATGCTGATGCTGATGCTGAT 3900
Db 5043 CTGAGCAGAAAGCCTTATGATGCTGATGCTGATGCTGATGCTGATGCTGATGCTGAT 5097

RESULT 5
US-09-866-866A-1
; Sequence 1, Application US/09866866A
; Patent No. US20020102244A1
; GENERAL INFORMATION:
; APPLICANT: Sorrentino, Brian
; APPLICANT: Schuetz, John
; TITLE OF INVENTION: A Method of Identifying and/or Isolating Stem Cells
; FILE REFERENCE: 1340-1-021CIP2
; CURRENT APPLICATION NUMBER: US/09/866, 866A
; PRIOR FILING DATE: 2001-08-30
; PRIOR APPLICATION NUMBER: 09/584, 586
; PRIOR FILING DATE: 2000-05-31
; PRIOR APPLICATION NUMBER: PCT/US99/11825
; PRIOR FILING DATE: 1999-05-27
; PRIOR APPLICATION NUMBER: 60/086, 988
; PRIOR FILING DATE: 1998-05-28
; NUMBER OF SEQ ID NOS: 27

; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 1
; LENGTH: 3860
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-866-866A-1
Query Match 74.8%; Score 3202; DB 10; Length 3860;
Best Local Similarity 89.6%; Pred. No. 0;
Matches 3455; Conservative 0; Mismatches 395; Indels 6; Gaps 1;
Qy 17 ATGATCTCTGAGGAGGCGCGTAAGGGAGTGCAGAGAACTCTGGAATATGGCAAA 76
Db 1 ATGATCTCTGAGGAGGCGCGTAAGGGAGTGCAGAGAACTCTGGAATATGGCAAA 60
Qy 77 AAAAAAAGTGAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAG 136
Db 61 AAAAAAAGTGAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAG 120
Qy 137 CGCTATTCATATGCTGATAGTGTATATGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTT 196
Db 121 CGCTATTCATATGCTGATAGTGTATATGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTT 180
Qy 197 CATGAGCTGCACTCCCTCTCATGATGCTGCTGTTTGAAGACATGACAGATAGTTTGA 256
Db 181 CATGAGCTGCACTCCCTCTCATGATGCTGCTGTTTGAAGACATGACAGATAGTTTGA 240
Qy 257 AATGAGGAATTTCAAGAAACAAACCTTTTCCAGTTATTAATTAATGAAGTTTACGAA 316
Db 241 AATGAGGAATTTT-----GAAAGATCTGATGTCAAACTACATTAATGAAGTAAATTC 294
Qy 317 AATGAGGAATTTTCAATCAACCATCTGAGAGAGAAATGACCGATTCATTAATTAAT 376
Db 295 AATGATACAGGCTCTTCAATCAATCAATGAGAGAGACATGACCGATTCATTAATTAAT 354
Qy 377 AGTGGATCGGTGCTGAGGCTGCTGCTGCTGCTTACATCCAGTTTCATTTGCTGCTG 436
Db 355 AGTGGATCGGTGCTGAGGCTGCTGCTGCTGCTTACATCCAGTTTCATTTGCTGCTG 414
Qy 437 GCAGCAGGAAGAGATTAATCAAAATTTGAAGAAATTTTTCATGCTATCATGCGACAG 496
Db 415 GCAGCAGGAAGAGATTAATCAAAATTTGAAGAAATTTTTCATGCTATTAATGCGACAG 474
Qy 497 GAGATTTGCTGCTTGAAGTGTGATGCTGCTGCTGCTTAAACCCGCTCAGACGAT 556
Db 475 GAGATTTGCTGCTTGAAGTGTGATGCTGCTGCTTAAACCCGCTCAGACGAT 534
Qy 557 GTCTCCAAATCAATGAAGAAATTTGCGACAAATTTGAATGCTTCTCACTCAATAGCA 616
Db 535 GTCTCCAAATCAATGAAGAAATTTGCGACAAATTTGAATGCTTCTCACTCAATAGCA 594
Qy 617 ACATTTTTCACCGCTTTATATGTTGCTTACAGTGTGTTGAAGCTTAACCTTTGTGAT 676
Db 595 ACATTTTTCACCGCTTTATATGTTGCTTACAGTGTGTTGAAGCTTAACCTTTGTGAT 654
Qy 677 TTGGCATCAGCGCTGTTCTTGAAGCTTCAAGCCGCTGCTGCTGCTGCTGCTGCTGCTGCT 736
Db 655 TTGGCATCAGCGCTGTTCTTGAAGCTTCAAGCCGCTGCTGCTGCTGCTGCTGCTGCTGCT 714
Qy 737 TTACTGATTAAGAACTCTTGGCTTGAAGCAAGCTGAGCAAGCTGAGCAAGCTGAGCA 796
Db 715 TTACTGATTAAGAACTCTTGGCTTGAAGCAAGCTGAGCAAGCTGAGCAAGCTGAGCA 774
Qy 797 GCAGCATCAGAACTGATGATGCTTTGAGAGCAAGAAAGAAAGAAAGAAAGAAAGAAAG 856
Db 775 GCAGCATCAGAACTGATGATGCTTTGAGAGCAAGAAAGAAAGAAAGAAAGAAAGAAAG 834
Qy 857 AAAAAATTAAGAAAGCTTAAGAAATTTGGAATTTGAAGCTTCAAGCTTCAAGCTTCA 916
Db 835 AAAAAATTAAGAAAGCTTAAGAAATTTGGAATTTGAAGCTTCAAGCTTCAAGCTTCA 894
Qy 917 ATGTCGCGCTTCTATGATCTATGATCAATCATGCTGCTTCTGCTGCTGCTGCTGCTGCT 976
Db 894 ATGTCGCGCTTCTATGATCTATGATCAATCATGCTGCTTCTGCTGCTGCTGCTGCTGCT

Db 895 ATAGGTGCTCTTCCTGCTGATCTATGACTTTATGCTTGCCCTTCTGCTATGGAGCC 954
Qy 977 TCCCTTGCTCTCTCCAGGAATATCTATTTGGACAAGTACTACTGTCCTTTCTGTGA 1036
Db 955 ACCCTTGCTCTCTCAGGGGAATATCTATTTGGACAAGTACTACTGTCCTTTCTGTGA 1014
Qy 1037 TTAATTTGGGGCTTTTACTATTGTGACAGGCAATCCCAAGCATTTGAAGATTTGGACAAGCA 1096
Db 1015 TTAATTTGGGGCTTTTACTATTGTGACAGGCAATCCCAAGCATTTGAAGATTTGGACAAGCA 1074
Qy 1097 AGAGAGAGAGCTTATGAATCTTCAAGATTAATGACATTAACCAAGCATTTGACAGCTAT 1156
Db 1075 AGAGAGAGAGCTTATGAATCTTCAAGATTAATGATTAATGATTAAGCCAGATTTGACAGCTAT 1134
Qy 1157 TCGAAGAGTGGGACCAACAGATTAATTTAAGGAAATTTGAAATTTGAAATTTGCAAC 1216
Db 1135 TCGAAGAGTGGGACCAACAGATTAATTTAAGGAAATTTGAAATTTGAAATTTGCAAC 1194
Qy 1217 TTTAGTTACCCCTTCTCGAAAAAGAAAGTTAAGATCTTAAAGGCTTCAACCTGGAAGTTGAG 1276
Db 1195 TTTAGTTACCCCTTCTCGAAAAAGAAAGTTAAGATCTTGAAGGCTTGAACCTGGAAGTTGAG 1254
Qy 1277 AGTGGGAGACAGTGGGCTGTTGGGAACAGTGGCTGGGGGAAAGACACGCTGCAG 1336
Db 1255 AGTGGGAGACAGTGGGCTGTTGGGAACAGTGGCTGGGGGAAAGACACACGCTGCAG 1314
Qy 1337 CTGATGAGAGGCTCTATGACCCACAGATGGCATGGTCTGTATTGATGAGACAGGACATT 1396
Db 1315 CTGATGAGAGGCTCTATGACCCACAGATGGCATGGTCTGTATTGATGAGACAGGACATT 1374
Qy 1397 AGGACCATTAATGTAGGATCTCTGGGAAATTTACTGGGTGGTGGTGGTGGTGGTGGTGG 1456
Db 1375 AGGACCATTAATGTAGGATCTCTGGGAAATTTACTGGGTGGTGGTGGTGGTGGTGGTGG 1434
Qy 1457 TTGTTTGGACACAGATTAAGGATTAAGGATTAAGGATTAAGGATTAAGGATTAAGGATTA 1516
Db 1435 TTGTTTGGACACAGATTAAGGATTAAGGATTAAGGATTAAGGATTAAGGATTAAGGATTA 1494
Qy 1517 GAGATTTGAGAAAGCTGTTAAGGAAGCCCAATGCTATGATTTTATCATGAAATCACTAAT 1576
Db 1495 GAGATTTGAGAAAGCTGTTAAGGAAGCCCAATGCTATGATTTTATCATGAAATCACTAAT 1554
Qy 1577 AAATTTGACACTCTGTTGGAG 1636
Db 1555 AAATTTGACACTCTGTTGGAG 1614
Qy 1637 ATCGGCATTTGCTGGGGCCCTGGTGGCAACCCCAAGATTTCTTCTGATGAGAGCAAG 1696
Db 1615 ATCGGCATTTGCTGGGGCCCTGGTGGCAACCCCAAGATTTCTTCTGATGAGAGCAAG 1674
Qy 1697 TCAGCTCGACACTGAAGAGTGAAGAGTTCAGGTGGCCCTGGATTAAGGCAAGAAAA 1756
Db 1675 TCAGCTCGACACTGAAGAGTGAAGAGTTCAGGTGGCCCTGGATTAAGGCAAGAAAA 1734
Qy 1757 GGGCGGACATCCATTTGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1816
Db 1735 GGGCGGACATCCATTTGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1794
Qy 1817 GCTGGTTTGAATGAGAGTTCATTTGAGAGAAAGAAATCATGATGATGATGATGATGATG 1876
Db 1795 GCTGGTTTGAATGAGAGTTCATTTGAGAGAAAGAAATCATGATGATGATGATGATGATG 1854
Qy 1877 AAGGCAATTTACTTCAAACTGTGCATATGCAAGACAAGAAAGAAATGAATGATGATGATG 1936
Db 1855 AAGGCAATTTACTTCAAACTGTGCATATGCAAGACAAGAAAGAAATGAATGATGATGATG 1914
Qy 1937 AATGGCACTGTGTAATCCAAAGTGAATGATGATGATGATGATGATGATGATGATGATG 1996
Db 1915 AATGGCACTGTGTAATCCAAAGTGAATGATGATGATGATGATGATGATGATGATGATG 1974
Qy 1997 GGGTCCAGTTTAT 2056
Db 1975 AGATTCAGTCTATAGAAAAAGATCAACTGCTAGAGTGTCCGTGATTCACAAGCCCAA 2034

Qy 2057 GACAAAAAGCTTGTACAAAAAGAGACTTGAATGAGATGATGATGATGATGATGATG 2116
Db 2035 GACAAAAAGCTTGTACAAAAAGAGAGCTGTGATGAAGATTAATGATGATGATGATGATG 2094
Qy 2117 AGGATTTGAGAGCTGAACTCAACAGTGAATGGCTTATTTTGTGGTGGTGGTGGTGGTGG 2176
Db 2095 AGGATTTGAGAGCTTGAATTAATTTTACTGAATGGCTTATTTTGTGGTGGTGGTGGTGG 2154
Qy 2177 ATTTAAGAGAGAGGCTGCAACAGATTTTCAATTAATTTTCAAGATTAATGAGATG 2236
Db 2155 ATTTAAGAGAGAGGCTGCAACAGATTTTCAATTAATTTTCAAGATTAATGAGATG 2214
Qy 2237 TTTACCCGAGATGGAGATGCTGGAACAAAAAGACAAGATTAATGATGATGATGATGATG 2296
Db 2215 TTTACCCGAGATGGAGATGCTGGAACAAAAAGACAAGATTAATGATGATGATGATGATG 2274
Qy 2297 TTTCTAGCTCTTGAATTAATTTCTTTTATTTATTTCTCTGACAGGCTTCAATTTGGC 2356
Db 2275 TTTCTAGCTCTTGAATTAATTTCTTTTATTTATTTCTCTGACAGGCTTCAATTTGGC 2334
Qy 2357 AAAGCTGGGAGATGCTGCAATGAGGCTTGCATACATGTTTTCAGATGCTGAG 2416
Db 2335 AAAGCTGGGAGATGCTGCAATGAGGCTTGCATACATGTTTTCAGATGCTGAG 2394
Qy 2417 CAGGATGTACGCTGTTTGTATGACCTTAATAAACACACAGTGGATGACACAGGCTT 2476
Db 2395 CAGGATGTACGCTGTTTGTATGACCTTAATAAACACACAGTGGATGACACAGGCTT 2454
Qy 2477 GCCAATGATGCGGCTCAATTTAAGGGCTATAGGCTTCCAGGCTTGCATACACAG 2536
Db 2455 GCCAATGATGCGGCTCAATTTAAGGGCTATAGGCTTCCAGGCTTGCATACACAG 2514
Qy 2537 AATATAGCAAACTTTGGGACAGCATTTATTTATCTTAATCTTAATGTTGGCAATTACA 2596
Db 2515 AATATAGCAAACTTTGGGACAGCATTTATTTATCTTAATCTTAATGTTGGCAATTACA 2574
Qy 2597 CTTTACCTCTTACGAACTTGTACCATCATTTGCAATTAAGAGAGTGTGAAATGAAATG 2656
Db 2575 CTTTACCTCTTACGAACTTGTACCATCATTTGCAATTAAGAGAGTGTGAAATGAAATG 2634
Qy 2657 TTGCTGGACAAGCACTGAAGATTAAGAAAGAGCTTAAGAGAGTGGGAATGCTGACA 2716
Db 2635 TTGCTGGACAAGCACTGAAGATTAAGAAAGAGCTTAAGAGAGTGGGAATGCTGACA 2694
Qy 2717 GAAGCATGCAAACTTCCGAACCTGTTCTTGACTGGGAGACGAAGTTGAAATAC 2776
Db 2695 GAAGCATGCAAACTTCCGAACCTGTTCTTGACTGGGAGACGAAGTTGAAATAC 2754
Qy 2777 ATGTATGACAGAGTTTGGCAAGTACATACAGAACTTTTGAAGAACACACATCTTC 2836
Db 2755 ATGTATGACAGAGTTTGGCAAGTACATACAGAACTTTTGAAGAACACACATCTTC 2814
Qy 2837 GGGGTCTCATTTTCTATACCCAGGCAATGATGATTTTCTCATGCTGGATGTTCCGG 2896
Db 2815 GGGGTCTCATTTTCTATACCCAGGCAATGATGATTTTCTCATGCTGGATGTTCCGG 2874
Qy 2897 TTTGGTCCCTACTTGGTGGCAATGATGATGATGATGATGATGATGATGATGATGATGATG 2956
Db 2875 TTTGGTCCCTACTTGGTGGCAATGATGATGATGATGATGATGATGATGATGATGATGATG 2934
Qy 2957 TCAGCTATGCTTTTGGTGGCAATGATGATGATGATGATGATGATGATGATGATGATGATG 3016
Db 2935 TCAGCTATGCTTTTGGTGGCAATGATGATGATGATGATGATGATGATGATGATGATGATG 2994
Qy 3017 GCCAAAGCCAAAGTATCCAGACACCCACAGTCAATGATGATGATGATGATGATGATGATG 3076
Db 2995 GCCAAAGCCAAAGTATCCAGACACCCACAGTCAATGATGATGATGATGATGATGATGATG 3054
Qy 3077 GACAGCTACAGCCCTCAGGCTCAAGCCCAATAGCTTGAAGGAATAGGACATTTAAT 3136
Db 3055 GACAGCTACAGCCCTCAGGCTCAAGCCCAATAGCTTGAAGGAATAGGACATTTGAT 3114

```

: LENGTH: 3860
: TYPE: DNA
: ORGANISM: homo sapiens
: OS-09-866-866A-3

Query Match      74.8%   Score 3200.4;   DB 10;   Length 3860;
Best Local Similarity 89.6%;   Pred. No. 0;
Matches 3454;   Conservative 0;   Mismatches 396;   Indels 6;   Gaps

OY 17 ATGATCTCTGAGAGAGCCGCTTAAGGGGAGCTGACAGAGAAACCTTCTGGAAATGGCGAA 76
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1  ATGATCTTGAAGAGGCGCCGCAATGAGAGAGCAAGAAAGAAACCTTTTAAACAGAAC 60

OY 77 AAAGTAAAAAATGAGAGAAAGAAAGAAACCAACACTGTGACGCTTGGCAATGTT 136
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 61 AATTAATGTAAAAAGATTAAGAAAGAAAGAAACCAACTGTGATGATTTTCAATGTT 120

OY 137 CGCTATTCAAAATGGCTTATAGTGTGTATATCTGGTGGGACAATGGCGCATATC 196
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 121 CGCTATTCAAAATGGCTTATAGTGTGTATATCTGGTGGGACAATGGCGCATATC 180

OY 197 CATTGAGCGTCACTCCCTCTCATGATGCTGTTTTGGAAACATGACAGATAGCTTCCA 256
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 181 CATGGGCGCTGGAAGCTTCCCTCATGATCTGTGTTGGAGAAATGACAGATATCTTSCA 240

OY 257 AATGACGAAATTTCAAGAAACAAACCTTTCCAGTTATATTAATGAAAGATTACGAC 316
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 241 AATGACGGAATTTT-----GAAAGATCTGATGTCMAACATCATATATGAGATATTC 294

OY 317 AATACACACAAATTTTATCAACCAATCTGGAGAGAGAAATGACCAAGTATGCTATATAC 376
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 295 AATGATACAGGGGTTCTCATGATCTGTGAGAGAAACATGACAGATATGCTATATATAC 354

OY 377 AGTGGAGTGGTGGCGGGGCGTGTGGTGGCTCTTACATCCAGGTTTCATTGTGGTCTG 436
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 355 AGTGGAGTGGTGGCGGGGCGTGTGGTGGCTCTTACATCCAGGTTTCATTGTGGTCTG 414

OY 437 GCAGCAGAGAAACAGATCTCAAAATTTGAAAACAATTTTTCATGCTATCAGCACAG 496
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 415 GCACCTGGAAGACAAATATACACAAAATTTGAAAACAGTTTTTTCATGCTATATGCGACAG 474

OY 497 GAGATTGGCGGTTTACGTGATGATGACGTGTGGGAGCTTAACACCGGCTCACAGACAT 556
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 475 GAGATTAGCTGTTTATATGACAGATGTGGGAGGCTTAACACCGGCTCACAGACAT 534

OY 557 GTCTCAAATCAATGAGGAATGTGGCGCAAAATTTGAATGTTCTTCTACTCAATAGCA 616
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 535 GTCTCAAAGTATATGAGGATTAATGCTGACAAAATTTGAAATGTTCTTCACTCAATAGCA 594

OY 617 ACATTTTCAACGGTTTTTATAGTGGGTTTTACAGCTGTTGGAAAGCTAACCTTGTGATT 676
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 595 ACATTTTCACTGGGTTTATAGTAGATTTTACACGCTGTTGAAAGCTAACCTTGTGATT 654

OY 677 TTGGCCATACACCGTGTCTTGGACCTTACAGCCGCAATGGGCAAAAGATCTATCTTCA 736
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 655 TTGGCCATACAGCTGTTCTTGGACCTTACAGCTGCTGTGGGCAAAAGATCTATCTTCA 714

OY 737 TTACTGATTAAGAACTCTTGCCCTATGCAAAAGCTGAGGACAGTATGCTGAAGAGCTTA 796
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 715 TTACTGATTAAGAACTCTTGCCCTATGCAAAAGCTGAGGACAGTATGCTGAAGAGCTTG 774

OY 797 GCAGCAATCAGAACTGATTTGCCCTTGGAGGACAAAGAAAGAACTTGAAGATGACAC 856
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 775 GCAGCAATCAGAACTGATTTGCCCTTGGAGGACAAAGAAAGAACTTGAAGATGACAC 834

OY 857 AAAAATTTAGAAAGAGCTTAAGGAATTTGGATTAAGAAAGCTATCCGCGCAACTTTCT 916
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 835 AAAAATTTAGAAAGAGCTTAAGGAATTTGGATTAAGGAAGAACTATTTACAGCAATTTCT 894

OY 917 AATGGGCGGCTTCTTATATGATCATGACATATCTCTGCGCTTTCTGTATGGAGAC 976
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 895 ATAGGCTGCTTCTCTGCTGATCTATCATCTTATCTCTGCGCTTCTGTATGGAGAC 954

```


D	2035	GACAGAAAGCTTAGTACCAGAAAGAGGCTCTGGATGAAAGTATACCTCAGTTTCCTTTTGG	2096
Q	2117	AGGATTTCGAAGCTGAACCTCAACGATGAGCTTATTTTGTGGTGTATATTTTGTCT	2176
D	2095	AGGATTATGAAGCTAAATTTTAACGATGAGCTTATTTTGTGGTGTATTTTGTCTC	2154
Q	2177	ATTATTAACGAGAGGCTCTGCAACGACATTTTTCATTAATATTTTTCACAGATTTATGAGATC	2236
D	2155	ATTATTAATGAGAGGCTCGCAACGACATTTGCAATTAATTTTTCAAAGATTATAGGGTT	2214
Q	2237	TTTACCCGAGATGAGATCCTGAAACAAACGACGAATAGTAAACATGTTTCTGTATG	2296
D	2215	TTTACGAAGATTTGATGATCTCTGAAACAAACGACGAATAGTAACTTTTTCATATG	2274
Q	2297	TTTCTAGTCTTGGAAATATTTCTTTTATTTACATTTTTCCTCCAGGCTTCACATTTGGC	2356
D	2275	TTTCTACCCCTTGGAATTTATTTCTTTTATTTACATTTTTCCTTCACAGGTTTACATTTGGC	2334
Q	2357	AAAGCTGGGAGATCCTCCTCAATGAGGCGTTCATTCATGTTTTCAGATCCATGCTAGA	2416
D	2335	AAAGCTGGAGATCTCTCCACAGGCGCTCGATACATAGTTTTCATCCATCCATGCTAGA	2394
Q	2417	CAGAGTCAAGCTGTTTGTATGACCCATAAAACACCACTGGAGATGACAAACAGGCTT	2476
D	2395	CAGAGTGAAGTTGTTGTATGATACCCATAAAACACCACTGGAGATGACATCCAGATC	2454
Q	2477	GCCAAATGAGGGGCTCAAGTTTAAAGGGCTATAGGTTCCAGGCTTGCCTATTACCCAG	2536
D	2455	GCCAAATGAGGCTCAAGTTTAAAGGGGCTATAGGTTCCAGGCTTGCCTATTAATACCCAG	2514
Q	2537	AATATAGCAAAATCTTGGGACAGGATTTATTTATCTTATCTATGATGTTGGCAATTACA	2596
D	2515	AATATAGCAAAATCTTGGGACAGGATTTATTTATCTTATCTATGATGTTGGCAATTACA	2574
Q	2597	CTTTTACCTTAGAATTTGATCCATATTTGCAATAGCAGAGATTTGAAATGAAATG	2656
D	2575	CTGTTACTCTTAGAATTTGATCCATATTTGCAATAGCAGAGATTTGAAATGAAATG	2634
Q	2657	TTTGTGACAAAGCTAGAAATAGAAAGAGCTGAAAGAGCTGGGAAAGTTGCTACA	2716
D	2635	TTTGTGTGACAAAGCTAGAAATAGAAAGAGCTGAAAGAGCTGGGAAAGTTGCTACT	2694
Q	2717	GAGCAATGCAAACTTCCCACTGTTGTTCTTTGACTGGGACAGAAAGTTGAATAC	2776
D	2695	GAGCAATGCAAACTTCCCACTGTTGTTCTTTGACTGGGACAGAAAGTTGAATAC	2754
Q	2777	ATGTATGCACAGATTTGCAAGTACCATTCAGAAACTCTTTGAGAAAGCAACATCTTC	2836
D	2755	ATGTATGCACAGATTTGCAAGTACCATTCAGAAACTCTTTGAGAAAGCAACATCTTC	2814
Q	2837	GGGGTCTCATTTTCTATCCACCAGCAATGATATTTTTCCTATGCTGATGTTTCCGG	2896
D	2815	GGAAATTCATTTTCTCTTCCACCAGCAATGATATTTTTCCTATGCTGATGTTTCCGG	2874
Q	2897	TTTGTGCTACTTGGTGGCAATAGTTCATGAACTTTCAGGATGTTCTTGGATTC	2956
D	2875	TTTGTGCTACTTGGTGGCAATAGTTCATGAACTTTCAGGATGTTCTTGGATTC	2934
Q	2957	TGAGCTATTGTTTGGTGCCATGGCAGTGGGAGGCTGATTCATTTGCTCCGCACTAT	3016
D	2935	TGAGCTATTGTTTGGTGCCATGGCAGTGGGAGGCTGATTCATTTGCTCCGCACTAT	2994
Q	3017	GCCAAAGCCAAAGTTTCAGCAGCCACAGCTCATATGATCTTAAAAAAGCCCTGATT	3076
D	2995	GCCAAAGCCAAAGTTTCAGCAGCCACAGCTCATATGATCTTAAAAAAGCCCTGATT	3054
Q	3077	GACACTACAGCCTCAGCGGCTCAAGCCAAATACGTTGGAAGAAATGACATTTAAT	3136
D	3055	GACACTACAGCAGGAAAGGCTTAATCGCAACACATTTGGAAGAAATGACATTTGAT	3114
Q	3137	GAGGTCGTTCACTATCCACTGACACAGACATCCCGTGTCTCCAGGCGTAGGCTTC	3196

Dp	3115	GAATTGTATATCACTATATCCACACCGAGACATCCAGATGCTTACAGGGACACTGACCTGG	31174
Qy	3197	GAGGTGAAGAAAGGGCCACAGCCCTGGCCCTCGTGGGTAGCAGTGGCTGTGGGAAAGACACNA	32565
Dp	3175	GAGGTGAAGAAAGGGCCACAGCCCTGGCCCTCGTGGGTAGCAGTGGCTGTGGGAAAGACACNA	32343
Qy	3257	GTTTGTACGTCTTGAAGCGCTTCTATGACCCCTTGGCTGGTTCAGTGTAAATGATGGC	33168
Dp	3235	GTGGTCCACGTCTCTGGAACGGGTTTTCAGACCCCTTGGCGAAGAAAGTGGCTGGTGGTGGC	32944
Qy	3317	AAAGAGATTAAGACACTGAATGTGCACGTGGCTCCGAGCACACCTGGGACATCGTGTCTAG	33767
Dp	3295	AAAGAAATTAAGCGACTGAATGTTCAGTGGCTCCGAGCACACCTGGGACATCGTGTCTAG	33544
Qy	3377	GAGCCCATCTGTTTGTACGTGAGCATTTGGCGGAGAAATTTGCTTGGAGACAAACACCGG	34366
Dp	3355	GAGCCCATCTCTGTTTGTACGTGAGCATTTGGCGGAGAAATTTGCTTGGAGACAAACACCGG	34144
Qy	3437	GTGGTATTCACATGAGACATATTATGACAGCGACAGAGAGGCGCAACATACACCATCTTATC	34866
Dp	3415	GTGGTATTCACAGGAAAGAAATGTGTAGGAGCAAGAGAGGCGCAACATACATACATCTTATC	34744
Qy	3497	GAGACACTCCCTTGAGAAATATCACACACAGATGAGACAAAGAAACCCAGCTCTCTGGT	35566
Dp	3475	GAGACACTCTCTTAAATTAATATGACACTAAAGATGAGACAAAGAAACCTCAGCTCTCTGGT	35344
Qy	3557	GGCCAGAAACACGCGATTTGCCATAGCTCGGCGCTCTTGTAGACACCTCATATTTTGGCT	36166
Dp	3535	GGCCAGAAACAAACGCGATTTGCCATAGCTCGGCGCTCTTGTAGACACCTCATATTTTGGCT	35944
Qy	3617	TTGGATGAAGCTACATCATCGCTCTGATGATAGAGAAAGTGGTTCACAAAGCCCG	36766
Dp	3595	TTGGATGAAGCCACATCATCGCTCTGATGATAGAGAAAGTGGTTCACAAAGCCCG	36544
Qy	3677	GACAAAGCCAGAGAGCGCCGACCTGGATTTGGATGCGCCACCGCTGTGCACATCATCG	37366
Dp	3655	GACAAAGCCAGAGAGCGCCGACCTGGATTTGGATGCGCCACCGCTGTGCACATCATCG	37144
Qy	3737	AATGCGAGTTTAATAGTGTGTTTCAGATATGCAAGTCAAGAGACATGGCACACATCA	37966
Dp	3715	AATGCGAGTTTAATAGTGTGTTTCAGATATGCAAGTCAAGAGACATGGCACACATCA	37744
Qy	3797	CACGTCTGTGCGCCAGAAAGGATCTATTTTTCATGTGTCACTGTCCAGGTGGAGACAAG	38566
Dp	3775	CACGTCTGTGCGCCAGAAAGGATCTATTTTTCATGTGTCACTGTCCAGGTGGAGACAAG	38344
Qy	3857	CGCTAGTAGAAGCTGTGG 3872	
Dp	3835	CGCTAGTAGAAGCTGTGG 3850	
RESULT 7			
US-09-866-866A-7			
Sequence 7, Application US/0986686A			
Patent No. US2002010224A1			
GENERAL INFORMATION:			
APPLICANT: Sorrentino, Brian			
TITLE OF INVENTION: A Method of Identifying and/or Isolating Stem Cells			
FILE REFERENCE: 1340-1-021CIP2			
CURRENT APPLICATION NUMBER: US/09/866, 866A			
PRIOR APPLICATION NUMBER: 2001-08-30			
PRIOR FILING DATE: 2000-05-31			
PRIOR APPLICATION NUMBER: PCT/US99/11825			
PRIOR FILING DATE: 1999-05-27			
PRIOR APPLICATION NUMBER: 60/086, 988			
PRIOR FILING DATE: 1998-05-28			
NUMBER OF SEQ ID NOS: 27			
SOFTWARE: PatentIn version 3.0			
SEQ ID NO 7			
TYPE: DNA			
LENGTH: 4788			

ORGANISM: Mus musculus
US-09-866-866A-7

Query Match	68.6%;	Score 2936.2;	DB 10;	Length 4788;
Best Local Similarity	84.4%;	Pred. No. 0;		
Matches 3334;	Conservative	0;	Mismatches 596;	Indels 19;
				Gaps 3;

QY	17	ATGATATCTGGAAGGAGGCCGTTAAGGGAGTGCAGAAAGAACATCTCTGGAAATATGGGCAA	76
Db	1	ATGGAACCTTGAAGAGACCCTTAAGGAAAGAGCAGACAAAGACTTCTCAAGATGGGCAA	60
QY	77	AAAAATGAAAAAATGAGAAAGAAAGAAACCAACGTCAGAGCATGTTTGCATGTTT	136
Db	61	AAGAGT---AAAAAGAGAAAGAAAGAAAGAAACCAAGCAGTCAGTCTTACAAATGTTT	117
QY	137	CGCTATTTCAATTTGGCTTGATAGTGTATATGTTGTTGGGACAAATGCTGCCATATC	196
Db	118	CGTTATGACAGTTTGGCTAGACAGGTTTATACATGCTGGTGGAAACCTGGCGGTATTATAC	177
QY	197	CATGAGACGTCACCTCCCTCATGATGCTGTTTTTGGAAACATGACAGATAGCTTTGCA	256
Db	178	CATGAGATGGGCTCCCACTTATGATCTGATCTTTGGTGAATAGACAGATGCTTTGCA	237
QY	257	AATGCGAAATTTTCAAGAAACAAACCTTTCCAGTATTAATTAATGAAGATATTAGAAC	316
Db	238	AGGTGAGAAAGCTCTTAATA-----CAGTACTAATATGATGAGGCC	282
QY	317	AATACACACATTTTCATCACCATCTGAGAGAGAAATATGACAGTATGCTATTATAC	376
Db	283	GATTAAGAAAGCCATGTTTGGCCAAACCTGAGAGAAAGAAATAGACCGTACGGCTACTATTAC	342
QY	377	AGTGGAGATGATGGCGGTGCGTGGTGGGCTGCTACATCCAGTTCATCTGGTCCG	436
Db	343	ACCGGATTTGGTGGTGGTGGCTCATAGTGGCTCATATCCAGTTCATTTTGGTCCG	402
QY	437	GCAGCGAAGACAAATCTCAAAATTTAGAAACAAATTTTTCATGCTATCAGACAG	496
Db	403	GCAGCTGGAAGACAGATACACAGATCAGGCAAGATTTTTCATGCTATTAATGATCAG	462
QY	497	GAGATGGCTGTTTGACGTGCATGAAGCTTGGGAGCTTAACCCGCGCTACAGACAT	556
Db	463	GAGATAGGCTGGTTTGAATGTGCATGAAGCTTGGGAGCTTAACCCGCGCTACAGATGAT	522
QY	557	GTCATCAAAATCAATGAAGATATGCGACAAAAATTTGGAATGTTCTTCACTCAATAGCA	616
Db	523	GTTTCCAAATTAATGAAGAAATTTGGTGAACAAATCGGAATGTTCTTCCAGGCAATGGCA	582
QY	617	ACATTTTACCCGTTTATATATGAGGGTTTACACGTGTTGGAAGCTAACCCCTTGTGAT	676
Db	583	ACATTTTGTGGTTTATATATGAGATTTTACCCGTGCTGGAAGCTTAACCCCTTGTGAT	642
QY	677	TTGGCCATCAGCCCTGTTCTTTGACCTTTCAGCGCCATCTGGGCAAGATACATATCTTCA	736
Db	643	TTGGCCATCAGCCCTGTTCTTTGACCTTTCAGCTGCTATTTGGGCAAGATATGCTTCA	702
QY	737	TTTACGATTAAGAAACCTTGGCGCTTATGAAAGCTGGAGAGTAGCTGAAGAAAGCTTCA	796
Db	703	TTTACGATTAAGAAACCTTTCATGCAAAAGCTGGAGACAGTCTGCTGAAGAAAGCTTCA	762
QY	797	GCAGCACTAGAACTGTGATGTGCTTTGAGAGACAAAGAAAGAACTTGAAGGATACAC	856
Db	763	GCAGCACTAGAACTGTGATGTGCTTTGAGAGACAAAGAAAGAACTTGAAGGATACAC	822
QY	857	AAAAATTTAGAAAGCAAGCTTAAGAAATTTGGATTAAGAAAGCAATACAGCGCAATCTT	916
Db	823	AAAACTTTGAAGAAAGCTTAAGAAAGCTTGGGATTAAGAAAGCAATACAGCGCAATCTC	882
QY	917	ATTGAGTGGCGTCTTATATGATCTATGATATATATGCTGTGGCTTCTGATATGGAGAC	976
Db	883	ATTGAGTGGCGTCTTCTCTATCTATGATATATATGCTGTGGCATCTGATATGGAGAT	942
QY	977	TCTTGTGCTCTCTCAGTGAATATTTATTTGACAAAGTACTCATGTCTCTTCTCTGA	1036

.Db	943	CCCTGGTCATCTCCAAAAGATACTCTATTGGACAAAGTGTCACTGTCTCTTTCCGTG	1002
QY	1037	TTAATTGGGGGCTTTTAGTATTGGACAGGATCCCAACGATTGAAGCATTTTGCACAAAGCA	1096
Db	1003	TTTAATTGGAGCANTCAGTGTGTGGACAGGCGATCTCCAAATATTGAAGCCCTTCGCAATGCA	1062
QY	1097	AGAGGAGCAGCTTATTGAAATCTTCAAGATTATGCAATTAACCAAGCATTGACAGCTAT	1156
Db	1063	CGAGGAGAGGCTTATGAAGTCTTCAAAATTAATTGATTATAGGCCAGTATAGCAGCTTC	1122
QY	1157	TCGAAGAGGAGCATTAACACAGATTAATTTAAGGAAATTTTGAATTCAAAATTTGTTCAC	1216
Db	1123	TCAAAGAGTGGGCAACAAACAGACACTTACAGAGAAATCTGGAAATTTAAGAAATATTTCAC	1182
QY	1217	TTCACTTACCTTCTCGAAAAGATTAGATCTTTAAAGGGTCTCAACCTGAAGGTTCAG	1276
Db	1183	TTTCACTTACCCATCTCGAAAAGATTAGATCTTTGAAGGGCTCTCAATCTGAAAGGTGAAG	1242
QY	1277	AGTGGGAGAGAGTGGGCGCTGGTTGGGACAGTGTGCTGGCGGAAGACAGACCGCTGACG	1336
Db	1243	ACGGGACAGAGGTGTGCCCCGTGGTGGCAACAGTGGCTGTGAAAAGAACAACTGTCCAG	1302
QY	1337	CTGATGCAAGAGCTCTATGACCCCAAGATGGCATGTCTGTATTGATGGACAGGACATT	1396
Db	1303	CTGAATGCAAAAGCTCTCAAGACCCCTAGATGGCATGTGATGCAACGACAGGACATC	1362
QY	1397	AGSACATTAATTAAGGACATCTTCGGGAAATTAAGTGTGTGTGTGAGTCAGAGAGCCGTGTG	1456
Db	1363	AGAACCATCATGTGTAGGTATCTGAGGGAATATTGTTGGTGTGTGTGAGTCAGAACCTGTG	1422
QY	1457	TTTGTTTGGCACCCAGATAGCTGAAAAACATTCGCTATGCGCCGGGAAATTTCCACATGAT	1516
Db	1423	CTGTTTTCCCAACGATGCGCCGAGAAACATTCGCTATGCGCCGGAAGATGTCCACCTGAT	1482
QY	1517	GAGATTGAGAAGCTGTTAAGGAAGCAATGCTATGATTTATCATGAACTACTAAT	1576
Db	1483	GAGATTGAGAAGCTGTCAAGGAAGCCAAATGCTATGACTTATCATATAAACTGCCCCAC	1542
QY	1577	AAATTTACACTGTGTTGGAGAGAGAGGGGCCACAGCTAGTGTGTGGACAGAAACAGAGA	1636
Db	1543	CAATTTGACACCTGTGTGTGTGGAGAGAGGGGCGACGTGATGTGGGGGACAGAAACAGAGA	1602
QY	1637	ATCGCACTTGTGGGGCCGTGGTTGGGCAACCCCAAGATTTCTGTCTGATGAAGCAACG	1696
Db	1603	ATCGCACTTGTGGGGCCGTGGTTGGGCAATCCCAAGATCTTGTGTGGAGAGGCAACC	1662
QY	1697	TCAGCTCTGACACTGAAAAGTGAAGCAGTGTCTCAGTGTGGCCCTGGAAATAGGCCAGAAA	1756
Db	1663	TCAGCCCTGGAATACAGAAAAGTGAAGCTGTGTGACGGCCGCACTGTGATAGGGCTAGAGAA	1722
QY	1757	GGCGGACATACACTGTGTATGACTATGCTTGTCTTACAGTTTGTGTAATGGCATGTCTATT	1816
Db	1723	GGCGGACACCACTGTGTATGACTATGCTATGCTTGTCTTACAGTTTGTGTAATGGCATGTCTATT	1782
QY	1817	GCTGTGTGTGATGATGAGTCAATGTGTGGAGAAAGAAATATATGATGAATCAATGAAGAAG	1876
Db	1783	GCTGTGTGTGATGATGAGTCAATGTGTGGAGAAAGAAATATATGATGAATCAATGAAGAAG	1842
QY	1877	AAGGGCATTTACTTAAACTGTGCACATGTGCAGCAAGAGGAAATGAATTGATTAGAA	1936
Db	1843	AAGGGCATTTACTTAAACTGTGTATGACACAGCACAGAGAAATGAATTGATTAGAA	1902
QY	1937	AATGCCACTGTGATCCAAAGTGAAGAGTATGCCCTTGGAAATGTCTCCAAAAGTTTCA	1996
Db	1903	AATGAAGCTTGTAAATGTAAAGATGAATTTGATTAATTTAGACATGTCTTCAAAAAGTTTCA	1962
QY	1997	GGGTCCACTTTAATTAAGAAAGATCACTGCGAGAGTATATACATGACACAGAGCCAA	2056
Db	1963	GGATCCACTTTAATTAAGAAAGATCACTGCGAAAGCATCTGTGACACATATACCAA	2022
QY	2057	GACAGAAAGCTTGTAGAAAAGAGACTTGAATGAAGAAATGTCACTCACTTCTTCTTGG	2116
Db	2023	GACAGAGAGCTTGTAGCAAAAGAGCCCTGTGATGAAGATGTCACTCACTTCTTCTTGG	2082

QY	2117	AGGATTCGAGCTGAACCTCACTCAAGATGAGCCTTATTTCTGGTGGTATTTATTTGCT	217
Db	2083	CGGATTCCTGAAGTTGAATTCACCTGAATGGCCATTATTTGTGGTATTTCTGTGC	214
QY	2177	ATTATAAGGAGGCCCGCAACCAAGATTTTCATAATATTTTCACAGATTATATGAGATC	223
Db	2143	ATATATTAATGAGAGCTTACAGCCAGAGATCTCCGTAATATTTTCATAAAGTTGAGGGATT	220
QY	2237	TTTACCAGATGAGATGCTGTGAATAACAAAGACAAATAGTATACATGTTTCTGTATTC	229
Db	2203	TTTACAATGTGGCCCCCTGAAACCCAGCGGCAACAGACACTGTATTTCTGTTGTTG	226
QY	2297	TTTCTAGTCCCTTGGAAATTTATTTCTTTTATTTACATTTTTCCTCAGGGCTTCATTTGAC	235
Db	2263	TTTTCTATCTCTTGGGATCATCTTTCTTTCTTATTTACATTTTTCCTCAGGGCTTCATTTGGC	232
QY	2357	AAAGCTGGGAGATCCTCACTCAAGGGCGCTGCATACATGGTTTTGACATTCATGCTGAGA	241
Db	2323	AAAGCTGGAGATATCTCTCCACAGAGCACTCCGATACAAATGGTTTTCAATTCATGCTGAGA	238
QY	2417	CAGGATGTCAGCTGTTGGTATGACCCTAATAAACACCACTGGAGCAATTGACAAACGAGTT	247
Db	2383	CAGGATGTCAGCTGTTGGTATGACCCTAATAAACACCACTGGAGCACTGACCAACAGCTC	244
QY	2477	GCCATGATGAGCGGCTCAAGTTAAAGGGGCTATAGGTTCCAGGCTTGCTGTATTTACCCAG	253
Db	2443	GCCAAAGATGCTGCTCAAGTGAAGAGGCTACAGGGCTTAGGGCTGTGTATTTTCAG	250
QY	2537	AATATAGCAATCTTGGGACAGGCAATTAATATCTTTAATCTAATGTTGGCAATTATACA	259
Db	2503	AACATAGCAAACTTGGGACAGGAGATCATATATCCCAATCTATGGCTGGCAACTACA	256
QY	2597	CTTTTACTCTTAGCAATTGTCCCATCATATTCGCAATTACAGAGATGTTGTAATGAAATG	265
Db	2563	CTTTTACTCTTAGCAATTGTACCCATATATTCGCAATTACAGAGATGTTGTAATGAAATG	262
QY	2657	TTGTCTGGACAGCACTGAAAGATTAAGAAAGAGTAAAGAGCTGGAGAGATTGGCTACA	271
Db	2623	TTGTCTGGACAGCACTGAAAGATTAAGAAAGAGTAAAGAGCTGGAGAGATTGGCTACA	268
QY	2717	GAACCCATCGAAACTCCCAACTGTTGTTCTTTGACTCGGAGACAGAAATTGAAATAC	276
Db	2683	GAACCAATTGAAACTCCCACTGTTGTTCTTTGACTCGGAGACAGAAATTGAAATAC	272
QY	2777	ATGTATGACAGAGTTTGCAGTACATACAGAAACTCTTGGAGAAACACACATCTTC	283
Db	2743	ATGTATGCCAGAGCTTGCAGATATACATACAGAAATGCCATAGAGAAACACACATCTTC	280
QY	2837	GGGGTCTCATTTTCTATCAACCCAGGCAATGATGATTTTTCTATAGCTGGCTTTCCGG	289
Db	2803	GGGATCACTGTTCTCTTCAACCCAGGCAATGATGATTTTTCTATAGCTGGCTTTCCGG	286
QY	2897	TTTTGTGCTACTTGTGTGCAATGATGATCAAGAACTTCAAGAGTCTTCTTGGTATTC	296
Db	2863	TTTTGTGCTACTTGTGTGCAACACACATCATGACTTTTGAATAAATGCTGTGTAATATC	292
QY	2957	TCGCTATATGCTTGTGTGCCATGGCAGTGGGGACGTCAGTTCTATTTGCTCTGACTAT	301
Db	2923	TCGCTATATGCTTGTGTGCCATGGCAGTGGGGACGTCAGTTCTATTTGCTCTGACTAT	298
QY	3017	GCCAAAGCCAAAGTATACAGCACGCCCAAGCATCATGATCATTTGAAAAAGCCCTCTGATT	307
Db	2983	GCGAAAGCAAGTGTCAAGCATCCCACTCATCAAGATCATTTGAAAAAGCCCGAGATTT	304
QY	3077	GACAGCTACAGCCCTCAAGGCTCAAGCCAAATACGTTGAGAGAAATGTGACATTTTAAT	313
Db	3043	GACAGCTACAGACCCCAAGGCTTAAAGCCGAATATGTTGGAAGAAATGTGCAATTTAGT	310
QY	3137	GAGGCTGCTTCAACTATATCCAGTCCAGTCCACACATCCCGCTGCTCCAGGGGCTGAGCTTC	319
Db	3103	GGAGTCTGCTTCAACTATATCCAGGCTCCAGTCCAGTCCAGGGGCTGAGCTTC	316

Db	2290	TCAAAGATTCCKRGATTCAGTCTAATAAGAAAGATCACTCGCAAAAGCAATCCGTGG	2349
Qy	2045	CCACAAAGGCCAAGACAGAAAGCTTGTAACAAAGAGACTTGATGAGATGACCTGCA	2104
Db	2350	CCACATGATCAAGACGGGGAAGCTTAGCAACCAAGAGGCTCGATGACGACGACTCTCA	2405
Qy	2105	GTTTCCCTCGAGGATTTCTGAAGCTGAACCTCAACTAATGGCTTAATTTTGTTGGT	2164
Db	2410	GCTTCCTTTGGGGATCTGTGAAGTTGAATTCAACTGAAGGCTTAATTTTGTTGGT	2469
Qy	2165	ATAATTTTGCTATTATATAACGAGAGGCTGCAACGACATTTTTCATTAATATTTTCAAG	2224
Db	2470	GTAATTTTGCCATPATATAAGGAGGCTTGCAACGACATTTCTCAATATTTTTCAG	2529
Qy	2225	ATTATAGGATCTTTATCCGAGATGAGGATCTGAAACAAAGACAGATAGTAACATG	2284
Db	2530	GTTGTAGGGGTTTATACAAAAGAACCCCTGGAATTCAGGGCGAGAAAGCAACTTG	2589
Qy	2285	TTTTTCGATTTGTTTCTATAGTCTCTTGGAATTAATTTCTTAATTTTCCATTTTCCACAGGC	2344
Db	2590	TTTTCTTAATTTGTCCGATCTCTGGATCATCTCTTTCATTAACGTTTTCCTTAAGGC	2649
Qy	2345	TTTACATTTTGCAAAAGCTGGGAGATCCTCATAAGGGCTTGTATACATGGTTTTCAGA	2404
Db	2650	TTTACATTTTGCAAAAGCTGGGAGATCCTCACCAGAGACTCCGATACATGTGTTTCAA	2709
Qy	2405	TCCATGCTGAGACAGATGTCAGTGGTTGTGATGACCTTAAAAACACACTGAGACATG	2464
Db	2710	TCCATGCTGAGACAGACATGATGCTGGTTGTGATGACCTTAAAAACACAGAGGCTG	2769
Qy	2465	ACAACAGGCTTGCATATGATGCGGCTCAAGTTAAAGGGGCTATAGTTCCAGGCTTGC	2524
Db	2770	ACCACACAGGCTTGCATATGATGCGCTCAAGTGAAGGGGCTACAGGGCTTAGGCTTGC	2829
Qy	2525	GTCATTACCCAGAAATATAGCAAACTTTGGGACAGGCAATTATATATCTTAATCTATGT	2584
Db	2830	GTTATTACCCAGAAACATAGCAAACTTTGGGACAGGCAATCATATATCCCTATCTACGGC	2889
Qy	2585	TGSCAATTAACACTTTACTCTTAGCAATTTGACCATATGCAATGCAATGAGAGTTGT	2644
Db	2890	TGSCAATTAACACTTTACTCTTAGCAATTTGACCATATGCAATGCAATGAGAGTTGT	2949
Qy	2645	GAAATGAAAAAGTTGTCTGCAACAACACTGAAAGATTAAGAAAGAGCTAGAGAGCTGG	2704
Db	2950	GAAATGAAAAAGTTGTCTGCAACAAGGCTGAAAGATTAAGAAAGAGCTAGAGAGCTGG	3009
Qy	2705	AAGATTGCTACAGAGCATGCAAAACTTCCGACACTGGTTCTTGTAGCTCGGGAGCAG	2764
Db	3010	AAGATCGCTACAGAGCAATTGAAAACTTCCGACACTGTCTCTTTGTAGCTCGGGAGCAG	3069
Qy	2765	AAGTTTGAATCATGTATGCAAGAGATTGCAATGACATACAGAAACTCTTTGAGAAA	2824
Db	3070	AAGTTTGAATCATGTATGCAAGAGCTTGCAGATACATACAGAAAGCTTTGAGAAA	3129
Qy	2825	GCACACATCTTGGGGCTCATTTTCTATACCCAGGCAATGATTTTCTTACTATGCT	2884
Db	3130	GCAGACGCTTTGGGATCACTTTCTCTTCCACCCAGGCAATGATTTTCTTACTATGCT	3189
Qy	2885	GGCTGTTCCGGTTGGTGCTACTTGTGGCAATGAGTTCAATGAATCTTTCAGATGTT	2944
Db	3190	GCTTGTTCGGTTTATCTACTATTGGGCGACAGAACTCATGACATTTGAAAAATGTT	3249
Qy	2945	CTTTTGTATTTCCACATATGCTTGTGGCAATGAGTGGGGCGAGCTCATTTT	3004
Db	3250	CTGTATATTTCCACATATTTGTCTTGTGGCAATGAGTGGGGCGAGCTCATTTT	3309
Qy	3005	GCTTCTGACTATGCCAAAGCAATATACAGACGCCACGTCATGATGATGAAAA	3064
Db	3310	GCTTCTGACTATGCCAAAGCAATGCTCGGATCCACATCATGAGATCATTTGAAA	3369
Qy	3065	AGGCGCTGATGAGAGCTAGAGCCCTACAGGCTCAAGCCAATACGTTGGAGAGAA	3124

Dh 1090 GAAGAAGCTTAGCAGCCATCAGAACTGTGATTCGCTTTGAGGACAAAGAGAACTT 1149
Qy 845 GAAAGTACAAACAAATTTTGAAGAACCTTAAGAAATTTGGATTAAGAAAGCTATCACG 904
Dh 1150 GAAAGGTACAAATTAACAAATTTTGAAGAGCTTAAAGGCTTGGATTAAGAAAGCTATCACG 1209
Qy 905 GCCAACTTTCTATTGTTGTCGCCCTTTTCTATTGATCTATGCATCATATGCTGTGGCATTC 964
Dh 1210 GCCAACTTTCCATTTGGGTGTCAGCTTTTCTGCTTATCTATCATATGCTGTGGCATTC 1269
Qy 965 TGTATGGGACCTCTGTTGCTCTCCAGTGAATATTCTATTGAGAACGACTACTACTGTC 1024
Dh 1270 TGTATGGGACCTCTGTTGCTCTCCAGTGAATATTCTATTGAGAACGACTACTACTGTC 1329
Qy 1025 TTTCTTTCTGTATTATTTGGGGCTTTTGTATTGAGAACGAGCATCCCAAGCATTTGAGCA 1084
Dh 1330 TTTTCTTTCTGTATTATTTGGAGCATTCAGTGTGGGACGAGCATCTCAAAATTTGAGGCC 1389
Qy 1085 TTTGCAAGCAAGAGAGAGAGAGCTTATGAATCTTCAAGATTAATGACATTAACCAAGC 1144
Dh 1390 TTTGCCCATATGAGAGAGAGAGAGCTTATGAAGTCTTCAATATTAATGATTAAGCCAGT 1449
Qy 1145 ATTGACGCTATTTCGAAGATGACATTAACCAAGATTAATTAAGGAAATTTGCAATTC 1204
Dh 1450 ATAGACAGCTTCTCAAAAGAGTGGGCAAAACCCGACACATACAAAGGAAATTTGCAATTC 1509
Qy 1205 AAAAATGTCACCTTCAGTTACCTCTCTGAAAAAGATTAAAGATCTTAAGAGTCTCAAC 1264
Dh 1510 AAAAATTTTACCTTCAGTTACCTCTCTGAAAAAGATTAAAGATCTTAAGAGTCTCAAC 1569
Qy 1265 CTGAAGGTTTCAGAGTGGGAGAGAGAGTGGGCTGTTGGAACAGTGGCTGGGGAGAGAGC 1324
Dh 1570 CTGAAGGTTGAAGGGGGAGAGAGAGTGGGCTGTTGGAACAGTGGCTGGGGAGAGAGC 1629
Qy 1325 AGCAGCGTGCAGCTGATGTCAGAGGCTCTATGACCCACAGATGAGTGTATTGAT 1384
Dh 1630 ACACGCTGCAGCTCTGAGAGGCTCTACGACCCCATAGAGGGCGAGTGCATGATGCAC 1689
Qy 1385 GGACAGAGATTAAGACCATTAATGTAAAGCATCTCGGGAATTTACTGCTGTGTAGT 1444
Dh 1690 GGACAGAGATCATGACCATCATATGTAGGTATCTGGGGAAATCATTTGGGTGTAGT 1749
Qy 1445 CAGGAGCCTGTGTGTTTGGCCACACGATAGCTGAAAAAATTCGCTATGAGCGCGGAAAT 1504
Dh 1750 CAGGAGCCCGTGTGTGTTGGCCACCAATTCGCGAATAATTCGCTATGAGCGCGGAAAT 1809
Qy 1505 GTACACATGATGAGATTGAGAAAGCTGTAAGAAAGCATGCTATGATTTTATCATG 1564
Dh 1810 GTACACATGATGATGATGAGAAAGCTGTAAGAAAGCATGCTATGATTTTATCATG 1869
Qy 1565 AAATCTACTAATTAATTTGACACTGTGTGAGAGAGAGAGGCGCCAGAGTGTGTGGA 1624
Dh 1870 AAATCTGCCCCCAAAATTTGACACCTGTGTGTGAGAGAGAGGCGCGAGCTGAGTGTGGA 1929
Qy 1625 CAGAAAGAGAAATGCGCATTTGCTCGGGCCCTGTTGCCAACCCCAAGATTTCTGTGCTG 1684
Dh 1930 CAGAAAGAGAGATGCGCATTTGCTCGGGCCCTGTTGCCAACCCCAAGATTTCTGTGCTG 1989
Qy 1685 GATAGGACAGCTGAGCTGTGACACTGAAAGTGAAGAGAGTGGTTCAGTGTGGCCCTGAT 1744
Dh 1990 GATAGGCGCAGCTGAGCTGTGACACAGAAAGAGGCGGTGTTGAGGCGCGCTGTGAT 2049
Qy 1745 AAGGCCGAGAAAGGCGGCACTACATTTGATGATGCTCATGTTTCTACAGTTGCTGAT 1804
Dh 2050 AAGGCTTGAAGAAAGGCGGAGCACGATTTGATGATGCTCACCGCTGTCTACAGTTGCTGAT 2109
Qy 1805 GCCGATGCTATTCCTGTTTGTATGATGAGATGATTTGGAGAAAGAAATCATGATGAA 1864
Dh 2110 GCTGAGCTATTCCTGTTTGTATGATGATGATTTGGAGAAAGAAATCATGATGAG 2169
Qy 1865 CTGATGAAGAGAGAGGCTTACTTCAACTGTGCACATGACAGACAAAGAGAAATGAA 1924
Dh 2170 CTGATGAGAGAGAGAAATTTACTTCAACTGTGCATGACTCAGACAGACGAGAAATGAA 2229

Qy 1925 ATTGATTAAGAAATGCCACTGTGTGATCCAAAAGTGAAGTGAATGCTTGGAAATGCT 1984
Dh 2230 ATTGATTAAGAAATGAAGCTTGTGAATCTAAACGGAATGTATGTGAGCATGCTCT 2289
Qy 1985 CCAAAAGATTAAGGATGAGTTTAATTAATTAAGAAAGATCAACCTGCGAGAGTATACATGCA 2044
Dh 2290 TCAAAAGATTTGGGATCCAGTCTAATTAAGAAAGATCAACCTGCGAGAGTATACATGCA 2349
Qy 2045 CCAACAGGCCAAGACAGAAAGCTTGTGTAACAAAGAGAGTGAATGAATGAGTACCTGCA 2104
Dh 2350 CCACATGATGACAGCGGGGAATTAAGCAACCAAGAGCTGTGATGACAGTACCTGCA 2409
Qy 2105 GTTTCCTTTGAGAGATTTGAAGCTGAATCTCAACATGAATGCTTATTTGTGTGCT 2164
Dh 2410 GCTTCCTTTTGGCGGATCTGAAAGTGAATCAACGAAATGCGCTTATTTGTGTGCT 2469
Qy 2165 ATATTTTGTCTATTATTAACGAGGCTGCAACCGATTTTCAATTAATATTTCAGG 2224
Dh 2470 GTATTTTGTCCATTAATTAATGAGGCTTGCAACCGATTTCTCATTAATATTTCAGG 2529
Qy 2225 ATATAGGATCTTTTCCCGAGATGAGATCTGAAACAAAGAGAGATGATACATG 2284
Dh 2530 GTTGTAGGGCTTTTACAAAAATGACACCCCTGAAATTCACCGGCAAGACGAACTTG 2589
Qy 2285 TTTTCTGTATTGTTTCTAGTCTTGAATTTATTTCTTTATTAATTTTCTTCCAGG 2344
Dh 2590 TTTTCTTTATTTGTTCTGATCTTGGATCATCTTTTATTTTCTTCTTCTTCCAGG 2649
Qy 2345 TTTACATTTGGGAAAGCTTGGGAGATCTCTACATAAGGCGCTTCATATGATTTTCA 2404
Dh 2650 TTTACATTTGGGAAAGCTTGGGAGATCTCTACATAAGGCGCTTCATATGATTTTCA 2709
Qy 2405 TCCATGCTGAGACAGATATGATGCTGTTGATGACCTTAACCAACACATGAGAGCTTG 2464
Dh 2710 TCCATGCTGAGACAGATATGATGCTGTTGATGACCTTAACCAACACATGAGAGCTTG 2769
Qy 2465 ACAACAGGCTTCCCAATGATGCGGCTCAAGTTAAAGGCGCTATAGTTCAGAGCTTGT 2524
Dh 2770 ACCACAGGCTTCCCAATGATGCGGCTCAAGTTAAAGGCGCTATAGTTCAGAGCTTGT 2829
Qy 2525 GTCATTTACCAAAATATAGCAAAATCTTGGGACAGCAATTAATATCTTAATCTATG 2584
Dh 2830 GTTATTTACCAAAATATAGCAAAATCTTGGGACAGCAATTAATATCTTAATCTATG 2889
Qy 2585 TGGCAATTAACATTTTACTTGAATGCAATTTGATCCATTCATTCAGAGAGTGTGT 2644
Dh 2890 TGGCAATTAACATTTTACTTGAATGCAATTTGATCCATTCATTCAGAGAGTGTGT 2949
Qy 2645 GAAATGAAATGTTGTGTCGACAGACGACGAAAGATTAAGAAAGCTTAAGAGAGCTG 2704
Dh 2950 GAAATGAAATGTTGTGTCGACAGACGACGTAAGAAAGATTAAGAAAGCTTAAGAGAGCTG 3009
Qy 2705 AAGATTCCTACAGAGAGCATGAAAACTTCGAACTGTTGTTCTTGAATGAGAGAG 2764
Dh 3010 AAGATTCCTACAGAGAGCATGAAAACTTCGAACTGTTGTTCTTGAATGAGAGAG 3069
Qy 2765 AAGTTTAATATGATGATGACAGAGTGTGCAAGTACCAATTAAGAAATCTTTGAGAA 2824
Dh 3070 AAGTTTAATATGATGATGACAGAGTGTGCAAGTACCAATTAAGAAATCTTTGAGAA 3129
Qy 2825 GCACACATCTTGGGGGTGCATTTTCTATCACCCAGGCAATGATTTTTCATATGCT 2884
Dh 3130 GCACACATCTTGGGGGTGCATTTTCTATCACCCAGGCAATGATTTTTCATATGCT 3189
Qy 2885 GCGTCTTCCGGTTGCTGCTTGTGATGATGATTTGATGATGATTTTTCATATGCT 2944
Dh 3190 GCGTCTTCCGGTTGCTGCTTGTGATGATGATTTGATGATGATTTTTCATATGCT 3249
Qy 2945 CTTTTGTATTTCTGAGATTTGCTTTTGTGATGATGATGATGATGATGATGATTTT 3004
Dh 3250 CTTTTGTATTTCTGAGATTTGCTTTTGTGATGATGATGATGATGATGATGATTTT 3309

3005 GCTCCGATATGCCAAGCAAGATATACAGACCCACGTCATCATGATCATTTGAAAA 3064
3310 GCTCTGATACAGCAAGCAAGCAAGTCTCGCATCCCATCATCATCATCATTTGAAAA 3369
3065 AGCCCTGATGACAGCTACAGCCCTCAAGCCGCAAGCAAGTACGTTGAGGAAGT 3124
3370 ATCCCTGATGATGACAGCTACAGCCGCAAGGAGGCTTTGAAAGCTTAATGTTGAAAGT 3429
3125 GTCACATTAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 3184
3430 GTCACATTAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 3489
3185 GGGCTGAGCTGAGGATGAGGAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 3244
3490 GGGCTGAGCTGAGGATGAGGAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 3549
3245 GGGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 3304
3550 GGGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 3609
3305 CTAAATGATGAGCAAGAGATGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 3364
3610 TTTCTGATGATGAGCAAGAGATGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 3669
3365 ATCTGATGATGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 3424
3670 ATCTGATGATGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 3729
3425 GACAAACAGCCGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 3484
3720 GACAAACAGCCGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 3789
3485 GACCACTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 3544
3790 GACCACTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 3849
3545 GACCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 3604
3850 GACCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 3909
3605 CATATTTGCTTTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 3664
3910 CATATTTGCTTTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 3969
3665 CAGAAGCCCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 3724
3970 CAGAAGCCCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 4029
3725 TCCACATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 3784
4030 TCCACATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 4089
3785 GGCACATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 3844
4090 GGCACATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 4149
3845 GCTGAGCAAGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 3904
4150 GCTGAGCAAGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 4208
3905 GCTGAGCAAGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 3957
4209 GCTGAGCAAGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 4261

RESULT 10
US-09-866A-5
; Sequence 5, Application US/09866866A
; Patent No. US20020102244A1
; GENERAL INFORMATION:
; APPLICANT: Sorrentino, Brian
; APPLICANT: Schmetz, John
; TITLE OF INVENTION: A Method of Identifying and/or Isolating Stem Cells

FILE REFERENCE: 1340-1-021CIP2
CURRENT APPLICATION NUMBER: US/09/866, 866A
CURRENT FILING DATE: 2001-08-30
PRIOR APPLICATION NUMBER: 09/584, 586
PRIOR FILING DATE: 2000-05-31
PRIOR APPLICATION NUMBER: PCT/US99/11825
PRIOR FILING DATE: 1999-05-27
PRIOR APPLICATION NUMBER: 60/086, 988
NUMBER OF SEQ ID NOS: 27
SOFTWARE: PatentIn version 3.0
SEQ ID NO: 5
LENGTH: 4189
TYPE: DNA
ORGANISM: Mus musculus
US-09-866-866A-5
Query Match 63.5%; Score 2716.6; DB 10; Length 4189;
Best Local Similarity 81.7%; Pred. No. 0;
Matches 3215; Conservative 0; Mismatches 694; Indels 25; Gaps 6;
17 ATGATCTGTAAGAGAGGCGGTAAAGGAGTGCAGAGAACTTCTGGAATGGGCAAA 76
1 ATGAGAGTTGAGAGAACTTAAAGGAGAGAGCAAGAACTTCTGGAATGGGCAAA 60
77 AAAAGTAAAAAATGAG 117
61 AAGAGT---AAAAG 117
137 CGCTATTCGAATGCTGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 196
118 CGCTATGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 177
197 CATGAGAGTGCATCCCTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 256
178 CATGAGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 237
257 AATGAGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 316
238 AAGAGAG-----AAGCCGATTTCTGCAAGCATTTCTATATCAAAAGTGAACCAAC 288
317 AATGAGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 373
289 AGTACTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 348
374 TACAGTGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 433
349 TACAGTGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 408
434 CTGGCAGCAG 493
409 CTGGCAGCAG 468
494 CAGAGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 553
469 CAGAGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 528
554 GAT 613
529 GAT 588
614 GCAAGATTTTCAACCGGTTTATAGTGGGTTTACACCGGTTTGAAGTAAACCTTGTG 673
589 ACCAGATTTTACCGGTTTATAGTGGGTTTATAGTGGGTTTGAAGTAAACCTTGTG 648
674 ATTTGGCAGTCAAGCCCTGTTTGGAGTTTCAAGCCCATCTGGGCAAGATTAATCT 733
649 ATTTGGCAGTCAAGCCCTGTTTGGAGTTTCAAGCCCATCTGGGCAAGATTAATCT 708
734 TCATTACTAGTAAG 793
709 TCATTACTAGTAAG 768

QY	794	TTTGACGACATCAGACATGTGATTTGCTTTGGAGGACCAAAAGAAGACTTGGAAAGTAC	853
Db	769	TTAGGACGACATCAGACATGTGATTTGCCTTTGGAGGACCAACAGAAAGAACTTGGAAAGTAC	828
QY	854	AAACAAAATTTGAGAAAGCTTAAAGCAATTGGATTAAGAAAGCAATACAGGCAACATT	913
Db	829	AAATAAAATTTTGAAGAAAGCTTAAAAATTTGGATTAAGAAAGCAATACAGGCAACATT	888
QY	914	TCATTTGGTGCGCTTCTCTATTGATCTATGCAATATGCTCTGGCTTCTGTGTGGG	973
Db	889	TCGATAGGACATTGGCTACGTGTGTGTCTATGCAATATGCAATGCGGATTTCTGTGTGGG	948
QY	974	ACCTCCCTGGGCTCTCCACATGAATTTCTATTGGACAAGTACTACTGTCTCTTTTCT	1033
Db	949	ACATCTCTGGTCCCTCAATTAAGAAATTTCTATTGGAGAAGTGTCTACTGTCTCTTCTCT	1006
QY	1034	GTATTAAATTGGGGCTTTTATTATTGGACGCGCATCCCAACATTTGAAAGCTTTGCAAC	1093
Db	1009	ATTTTGTGTGGGACTTTTACTATATGAGACCTTGGCCCCAACAATGAAAGCTTTGCAAC	1068
QY	1094	GCAAGAGGAGCAGCTTATGCAATCTTCAAGATTAATTGACATAAACCAAGCATTGACAC	1155
Db	1069	GCGAGAGGCGACCTTTGAAATCTTCAAGATTAATTGATTAACGACGCAAGCATTGACAC	1128
QY	1154	TATTGGAAGGTGACATTAACCAATATATTTTAAGGAAATTTGCAATTCAAAAATGTT	1213
Db	1129	TTCTCAACAAAGGGCTACAAACCAACAGTATAATGGGAAATTAAGATTTAAAAATGTT	1188
QY	1214	CACCTGATTAACCTTCTCTGAAAAGAACTTAAGATCTTTAAAGGGTCTCACTGTAAGGTT	1273
Db	1189	CACCTTCAACTACCCATTCAGTAAAGCAACAACTTGATAGATCTTGAAGGGCTCACTGTAAGGTT	1248
QY	1274	CAGAGTGGGACAGCATGTGGCGGTGGTTGGGAACAGTGGCGCGGGAAGGACACGACGCTG	1333
Db	1249	AAGAGGGGACAGACGGGTGGCCCTGGTTGGCAACAGTGGCGCTGGGGAAGGACCAACCTGTC	1308
QY	1334	CAGTGTATGCAAGGCTCTATGACCCACAGATGGCATGGTCTGTATTGATGACAGACAG	1393
Db	1309	CACCTATTCABAAGCTCTACGACCCCTGGAGGGCGGTGTGATGATGACAGCAAGAC	1368
QY	1394	ATTAGACCATTAATGTAAAGCATCTTTCGGGAATTAACGTGGTGGTATGACAGACCT	1453
Db	1369	ATCAGAACCATTAATGTAGGTATCTGAGGAGCATATTGGTGTGTGATGACAGAACT	1428
QY	1454	GTGTGTGTGGCCACACAGATTAAGCAAACTTCCGTATGCGCGCGGAATATGACCATG	1513
Db	1429	GTGTGTGTGGCCACACAGATTAAGCAAACTTCCGTATGCGCGCGGAATATGACCATG	1488
QY	1514	GATGAGATTGAAAGAGCTGTTAAGGAAGCAATGGCTATGATTTTATCATGAACACTACCT	1573
Db	1489	GATGAGATTGAAAGAGCTGTTCAAGGAAGCAATATGCTATGATTTTATCATGAACACTACCT	1548
QY	1574	AATTAATTTGCACTGTGTGTGGAGAGAGAGGGGCCAGCTGAGTGGTGGACGAAGAAACG	1633
Db	1549	CACCAATTTGACACCTGTGTGTGGAGAGAGGGGCCAGCTGAGTGGTGGACGAAGAAACG	1608
QY	1634	AGAAATCGCATTTGTGGGGCCCGGGTTGGCAACCCCAAGATTTCTGTCTGATGAGGCA	1693
Db	1609	AGAAATCGCATTTGTGGGGCCCGGGTTGGCGCAATCCCAAGATCTTTGTGTGACGAGGCC	1668
QY	1694	ACGTACGCTGGAACACTGAAGTAGAAGAGTGGTACAGTGGGCCCGGATTAAGGCGAGA	1753
Db	1669	ACCTCAGCCCTCGGATACAAAGATGAACCTGTGGTGAAGGCCCGCCTGATTAAGGCTGAGA	1728
QY	1754	AAAGCGCGGACATCAATGTGATTAACCTATGTTGTCTTCACTGTGTAATCCGATGCT	1813
Db	1729	GAAAGCGCGGACCAACATTTGTGATTAACCTATGTTGTCTTCACTGTGTAATCCGATGCT	1788
QY	1814	ATTTCGTGTTTGTATGATGAGCACTATTGTGAGAAAGGAATATGATGAATCATGTGAAA	1873
Db	1789	ATTTCGTGTTTGTATGATGAGTGTGCTATTGTGAGCAAGGAATATGATGAGCTCATGTGGA	1848
QY	1874	GAGAAAGGCAATTACTTCAAACTGTTCACAATGACAGCAAGGAATGAATTAAGTTA	1933

Db	1849	GAAGAAGGCACTTACTTCAAACTTGTCATGACACACBACACTAGAGGAAATGAACCA	1908
OY	1934	GAAAATGCCACTGTCATTCACAAAGCTGAAGTGAATGATGCTTGGAATGTCTCCAAAAGAT	19933
Db	1909	GGAATATATGCTTATGATATCCAGAGTGCACGTGATGCTTCGAACTGCATTCAGAAAGA	19688
OY	1994	TCAGGGTCCAGTTTAAATTAAGAAATCATCACTCCAGAGATATACATGCACACAGGC	20533
Db	1969	TCCAAATCACCTTAAAT---AAGGATATATTTACAGAAAGTGTCCACGAAACACAGAC	20255
OY	2054	CAGACACGAAGCTTGGTGCACAAAAGAGACTTGAATGGAATGTACCTCCAGTTTCTTC	21133
Db	2026	CAGAGACGAAGACTTATAGTATGAAAAGGCTGTGATGAAGATGTGGCTCTGCTTT	20855
OY	2114	TGGAGATTCCTGAAGCTGAATCAACTGAACTGAATGGCTTTTGTGGTTGGTATATTTGT	21733
Db	2086	TGGCGGATCCCTAAATCTAAATCTAATAGTAAATGGCTTATTTACTTGTGGGTGATCTTGC	21455
OY	2174	GCATATATAAGGAGGAGCCTGACACAGCATTTTCAATATATTTTTCAGGATTAAGG	22333
Db	2146	GCTGTTATTAATGGGTGCATCAACACATGTTTCCATAGATATTTTCAAGATTTGAGG	22055
OY	2234	ATCTTATCCGAGATGAGATTCCTGAACAACAGACAGATGTATCATGTTTCTGTGA	22933
Db	2206	GTTTTTTCAGAGATGATGACCATGAACATTAACAGACAGATGTATTTGTGTTCCGTG	22655
OY	2294	TGTTTTCAGTCCTTGGAAATATTTCTTTTATATACATTTTCTCCAGGGCTTCACATTT	23533
Db	2266	TTCCTTCTGGTTATGGGCTATATTTCTTTTGTATACATATTTCTTCAGGGCTTCACATTT	23255
OY	2354	GGCAAGCTGGGAGATCTCTACTAAGCGGCTTCGATACATGTGTTTCAGATCCAGTCG	24133
Db	2326	GGCAAGCGGAGAGATCTCTACCAACGAGTCCGATACATGTTTTCACAAATCCATGCTG	23855
OY	2414	AGACAGATGTACACTGTTTGTATGATGACCCCTAATAACACACTGGACATTTGAACACAG	24733
Db	2386	AGACAGGATTAATACCTGTGTTGATGACATTAACACACACTGGACATTTGAACACAG	24455
OY	2474	CTTCCCATGATGCGGCTCAAGTTTAAAGGGGCTATAGTTTCCAGGCTTCTCTCATTTAC	25333
Db	2446	CTCCCGAGTATGCTTCTACTGTTTAAAGGGGCGATGGCGCCAGGCTTCTGTACTTACC	25055
OY	2534	CAGATATATCAATCATCTGGGACAGGCAATTTATATCTTATATATGTTGGCAATTA	25933
Db	2506	CAGATATATCAAACTCGGAGACAGAGTCAATCTCTCTTATGTTAATGGCTGACGCTG	25655
OY	2594	ACACTTTTACTCTTAACCAATGTACCATCATTTGCAATAGCAGCATGTTGGAATGAA	26533
Db	2566	ACACTTCTACTCTTACTATTAATTACCGCTCATTTGATTTGGGGGAATTAATTGAATGAG	26255
OY	2654	ATGTTTCTGGAACACACTGGAAGATTAAGAAAGCGTATAGAGAGCTGGAGAGATTGCT	27133
Db	2626	CTGTCTCTGGCCAGACCTTGAAGACAAAGAAACGCTTGAGATCTCTGGGAAGATTGCT	26855
OY	2714	ACGAAAGCATGAAAACCTTCGAACTGTGTTCTTTGATCTCGGAGACAGAACTTTGAA	27733
Db	2686	ACAGAAAGCATTGAAGAACTTCGCGACTATGTGCTTTGATCTGGGAGACAGAACTTTGAA	27455
OY	2774	TACATGTATGACACAGTTTGGAAGTACCATACAGAAACATCTTTGAGGAAACACACATTC	28333
Db	2746	ACCATGTATGCCCAAGCTTGCAGGTACATACAGAAATCGATGAAGAAAGACACAGTG	28055
OY	2834	TTCGGGGTCTATTTTCTATACCCAGGACATGATGTATTTTCTATATGCTGCTCTTTC	28933
Db	2806	TTTGGGATTCAGTTTCTCTTCCACCAAGCCATGATGTATTTTCTTATGTGCTGTTTTC	28655
OY	2894	CGGTTTGGTGCCTACTTGGTGGCAAAAGAGTCAATGAACCTTTAGAGATGTCTTTGGTA	29533
Db	2866	CGGTTTGGTGCCTACTTGGTGGCAAAAGAGTCAATGAACCTTTAGAGATGTATGTGGTA	29255
OY	2954	TTCACAGTATTTGTTTGGTGGCCAGGACGTGGGGAGGTGAGTTCAATTTCTCTCTGAC	30133

QY 482 GCTATCATCGACAGAGATTGGCTGTTTGACGTCATGACGTTGGGAGCTTAACACC 541
DB 559 GCCATCATGAATCAGSAGATATAGCGTGGTTTACGCGTGAATGACGCTGGGAGCTCAACACC 618
QY 542 CGGCTCACAGACGATGTCTCCAAAATCAATAGAGAAATGGCGCAAAATTTGAAATGTC 601
DB 619 GCTTCACAGATGACGCTCCAAAATTAATAGAGAAATGGTGAACAACCTGGAAATGTTTC 678
QY 602 TTTCATCAATAGCAAACTTTTTCACCGGTTTATAGTGGGTTTACACGTGGTTGGAAG 661
DB 679 TTTGAGTCATAGACACTTTTTCACCGGTTTATTAATAGAGAAATTTAAAGTGGTTGGAAG 738
QY 662 CTAACCTTGATGATTGGCCATCAGCCCTGTTCTTGAGACTTACACGCGCATCTGSGCA 721
DB 739 CTAACCTTGATGATTGGCCCTGATGAGCCCTTATTTGGGTTGTCATCTGCGCATGSGCA 798
QY 722 AAGTACTATCTTACTTACTGATTAAGAACTCTTGCCCTATGCAAAAGCTGGAGACATTA 781
DB 799 AAGTACTAGCTTACTTACTTAAATAGAAAGCTCAGAGCTTATGCGAAAGCTGGAGAGTT 858
QY 782 GCTAAGAACTCTTACGACCAATCAGAACTGATGATGCTTTGGAGGACAAAAGAAAGA 841
DB 859 GCGCAAGAACTCTTACGACCAATCAGAACTGATGATGCTTTGGAGGACAAAAGAAAGA 918
QY 842 CTGAAAGGTACAAACAAAATTTAGAAAGCTTAAAGAAATTTGGATTAAGAAAGCTATTC 901
DB 919 CTGAAAGGTACAAATTAATTAAGAAAGCTTAAAGAAAGCTTGGCATTAAGAAAGCCATTC 978
QY 902 ACGGCCAACATTTCTATTTGCTCCGCTTCTTATGATCTATGCAATCATATGCTCTGCT 961
DB 979 ACGGCCAACATTTCTATTTGCTCCGCTTCTTATGATCTATGCAATCATATGCTCTGCT 1038
QY 962 TTCTGATGAGACCTCTCTGCTCTCTCAGTGAATTTCTATTTGAGCAAGTACTACT 1021
DB 1039 TTCTGATGAGACCTCTCTGCTCTCTCAGTGAATTTCTATTTGAGCAAGTACTACT 1098
QY 1022 GTCTCTTTTCTGATTAATTAATGAGGCTTTTATGATGAGCAAGGATCCCAAGCATTTGA 1081
DB 1099 GTCTCTTTTCTGATTAATTTATGAGGCTTTTATGATGAGCAAGGATCCCAAGCATTTGA 1158
QY 1082 GCATTTGCAAGGCAAGAGAGAGAGCTTATGAATTTCTCAAGATTAATGCAATTAACCA 1141
DB 1159 GCCTTTGCAAGGCAAGAGAGAGAGCTTATGAATTTCTCAAGATTAATGCAATTAACCA 1218
QY 1142 AGCATTTGACAGCTATTCGAAAGATGAGATTAACAGATTAATTAAGGAAATTTTGAA 1201
DB 1219 AGCATTTGACAGCTATTCGAAAGATGAGATTAACAGATTAATTAAGGAAATTTTGAA 1278
QY 1202 TTCAAAAATGTTCACTTCACTTACCTTCTCGAAAAGATTAAGATCTTAAAGGCTCTC 1261
DB 1279 TTTAAAATGTTTACTTCACTTACCTTCTCGAAAAGATTAAGATCTTAAAGGCTCTC 1338
QY 1262 AACCTGAAGTTTCAAGTGGGAGAGAGTGGGCTGGTTGGGAGAACGTGGCTGGGGAAG 1321
DB 1339 AACCTGAAGTTTCAAGTGGGAGAGAGTGGGCTGGTTGGGAGAACGTGGCTGGGGAAG 1398
QY 1322 AGCAGAACCTTGACGTGATGACAGAGCTCTATGACCCCAAGATGAGCATGGTCTGTATTT 1381
DB 1399 AGCAGAACCTTGACGTGATGACAGAGCTCTATGACCCCAAGATGAGCATGGTCTGTATTT 1458
QY 1382 GATGAGACAGCAATTAAGCAATTAAGCAATCTTCCGGAATTAAGTGGTGGTG 1441
DB 1459 GATGAGACAGCAATTAAGCAATTAAGCAATCTTCCGGAATTAAGTGGTGGTG 1518
QY 1442 AGTCAGAGAGCTGTGTTGTTTCCACCAAGATGAGTGAAGAAATTCGCTATGAGCGGAA 1501
DB 1519 AGTCAGAGAGAGCTGTGTTTCCACCAAGATGAGTGAAGAAATTCGCTATGAGCGGAA 1578
QY 1502 AATGTACCATGATGAGATTAAGAAAGCTGTTAAAGAGCAATGCTATGATTTATTC 1561
DB 1579 AATGTACCATGATGAGATTAAGAAAGCTGTTAAAGAGCAATGCTATGATTTATTC 1638
QY 1562 ATGAAGCTATCTAATTAATTTGACACTCTGGTTGAGAGAGAGGCGCCAGCTGAGTGT 1621

DB 1639 ATGAAGCTATCTAATTAATTTGACACTCTGGTTGAGAGAGAGGCGCCAGCTGAGTGTG 1698
QY 1622 GGACAGAAACAGAGATGCTTGGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1681
DB 1699 GGACAGAAACAGAGATGCTTGGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1758
QY 1682 CTGATGAGGCAAGCTCAGCTTGAGACATGAAGATGAAGAGAGAGAGAGAGAGAGAGAGAG 1741
DB 1759 TTGATGAGGCAAGCTCAGCTTGAGACATGAAGATGAAGAGAGAGAGAGAGAGAGAGAGAG 1818
QY 1742 GATAAGGCAAG 1801
DB 1819 GATAAGGCAAG 1878
QY 1802 AATGCCATGCTATTTGCTGTTTGAATGATGAGATGATGATGATGATGATGATGATGATGAT 1861
DB 1879 AATGCCATGCTATTTGCTGTTTGAATGATGATGATGATGATGATGATGATGATGATGATGAT 1938
QY 1862 GAACTCATGAAG 1921
DB 1939 GAACTCATGAAG 1998
QY 1922 GAAATTTGATTAAGAAATGCACTGATGATGATGATGATGATGATGATGATGATGATGATG 1961
DB 1999 GAAATTTGATTAAGAAATGCACTGATGATGATGATGATGATGATGATGATGATGATGATG 2058
QY 1982 TTCCCAAGAGATTTCAAGGCTTCAATTAATTAAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2041
DB 2059 ACTTCAGAAAGATTAAGAAATTTCAATTTAATTAAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2115
QY 2042 GCAACCAAG 2101
DB 2116 GCAACCAAG 2175
QY 2102 CCAAGTTCTCTGAGAGATTTCTGAGCTGAATCACTGAATGAGCTTATTTGAGTGT 2161
DB 2176 AAGTTCTCTCTGAGAGATTTCTGAGCTGAATCACTGAATGAGCTTATTTGAGTGT 2235
QY 2162 GGTATATTTTGTGATTAATTAAG 2221
DB 2236 GGTATATTTTGTGATTAATTAAG 2295
QY 2222 AGAATTAAGAGAGATTTTACCCGAGATGAGAGATCTGAAACCAACAGAGATTAAGTAC 2281
DB 2296 AGAATTAAGAGAGATTTTACCCGAGATGAGAGATCTGAAACCAACAGAGATTAAGTAC 2355
QY 2282 ATGTTTCTGATTAATTTCTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2341
DB 2356 ATGTTTCTGATTAATTTCTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2415
QY 2342 GGTTCATATTTGGCAAGAGTGGGAGATCTTCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2401
DB 2416 GGTTCATATTTGGCAAGAGTGGGAGATCTTCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2475
QY 2402 AGATTCATGCTGAGAGAGATGTCAGCTGTTGATGATGATGATGATGATGATGATGATGATGAT 2461
DB 2476 AGATTCATGCTGAGAGAGATGTCAGCTGTTGATGATGATGATGATGATGATGATGATGATGAT 2535
QY 2462 TTACACCAAGAGGTTGCCATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2521
DB 2536 TTACACCAAGAGGTTGCCATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2595
QY 2522 GCTGTCATTAACCAAGATTTAGCAATCTTGGGACAGAGATTAATTAATTAATTAATTAATTAAT 2575
DB 2596 GCTGTCATTAACCAAGATTTAGCAATCTTGGGACAGAGATTAATTAATTAATTAATTAATTAAT 2655
QY 2576 ATCTATGTTGGCAATTAACATTTTACTTGAATTTGATGATGATGATGATGATGATGATGATGAT 2635
DB 2656 ATCTATGTTGGCAATTAACATTTTACTTGAATTTGATGATGATGATGATGATGATGATGATGAT 2715
QY 2636 GGAGTTGTTGAATGAATGATGCTGGAGACAGACTGAAGATTAAGAAAGAGCTAGAA 2695

Db 2716 GGAATTATTGAATGAACCTGTTGCTGTCAGCCTTGAAGACAAGAAAGACCTAGAG 2775
OY 2696 GGAGCTGGAGAGATTGCTACAGAAAGCCATGAAAACCTTCCGAACCTGTTGTTTACT 2755
Db 2776 AACTCTGGGAGATGCTACAGAAAGCAATGAAAACCTTCCGACATGTTGCTGACT 2835
OY 2756 CGGAGACAGAGTTGAAATACATGATGACAGAGATTGCAAGTACATACAGAACTCT 2815
Db 2836 CGGAGACAGAGTTGAAATACATGATGACAGAGATTGCAAGTACATACAGAACTCT 2895
OY 2816 TTGAGAAAGACACATCTTCGGGCTCATCTTCTATCACCACAGCAATGATGATTTT 2875
Db 2896 TTGAGAAAGACACATCTTCGGGCTCATCTTCTATCACCACAGCAATGATGATTTT 2955
OY 2876 TCCATGCTGCTGTTTCCGGTTCGCTTCTACTTGGTGGCAAGAACTCATGACCTT 2935
Db 2956 TCCATGCTGCTGTTTCCGGTTCGCTTCTACTTGGTGGCAAGAACTCATGACCTT 3015
OY 2936 CAGAGATCTTCTTGGTATTCAGCTATTCCTTGGTGGCAAGTGGGAGGCTC 2995
Db 3016 GAAATATGTTATGTTGATTTCTGCTGTTCTGTTGTTGCTGCTGCTGCTGCTGCT 3075
OY 2996 AGTTCATTCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 3055
Db 3076 AGTTCATTCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 3135
OY 3056 AATGAAAAACCCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 3115
Db 3136 AATGAAAAACCCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 3195
OY 3116 GAAGAAATGATGATTTAATGATGATGATGATGATGATGATGATGATGATGATGATGAT 3175
Db 3196 GAAGAAATGATGATTTAATGATGATGATGATGATGATGATGATGATGATGATGATGAT 3255
OY 3176 GTGCTCAGGGGCTGAGCTTGGAGTGAAGAAAGGCGCAGAGCTGCTGCTGCTGCTGCTGCT 3235
Db 3256 GTGCTCAGGGGCTGAGCTTGGAGTGAAGAAAGGCGCAGAGCTGCTGCTGCTGCTGCTGCT 3315
OY 3236 AGTGGCTGGGAGAGACAGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 3295
Db 3316 AGTGGCTGGGAGAGACAGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 3375
OY 3296 GCTTCAGCTCAATTTGATGAGCAAGAAAGAACTGAAATGCTGCTGCTGCTGCTGCTGCT 3355
Db 3376 GGAACAGTGTCTTCTGATGAGCAAGAAATGAAAGCACTCAATGCTGCTGCTGCTGCTGCT 3435
OY 3356 CACCTGGGATCGTGTCTGAGAGAGCCATCTGTTGATGCTGAGATTCGCGAGCAATC 3415
Db 3436 ---CTGGGATGTGTCTGAGAGAGCCATCTGTTGATGCTGAGATTCGCGAGCAATC 3492
OY 3416 GCCTATGAGACAAAGCCGGGCTGATGATGAAAGATGATGATGATGATGATGATGATGAT 3475
Db 3493 GCCTATGAGACAAAGCCGGGCTGATGATGAAAGATGATGATGATGATGATGATGATGAT 3552
OY 3476 GCGAATACACACTTCTGAGACACTCCCTGAGAAATACACACAGATGATGATGATGATGAT 3535
Db 3553 GCGAATACACACTTCTGAGACACTCCCTGAGAAATACACACAGATGATGATGATGATGAT 3612
OY 3536 AAGGAAACCAAGCTCTCTGCTGAGAGAAACAGGATGATGATGATGATGATGATGATGAT 3595
Db 3613 AAGGAAACCAAGCTCTCTGCTGAGAGAAACAGGATGATGATGATGATGATGATGATGAT 3672
OY 3596 AGACAGCTCATTTTGTGTTGATGATGATGATGATGATGATGATGATGATGATGATGAT 3655
Db 3673 AGACAGCTCATTTTGTGTTGATGATGATGATGATGATGATGATGATGATGATGATGAT 3732
OY 3656 AAGGTTCTCAAGAGCCCTGAGCAAGCAAGAGAGCCGACGATGATGATGATGATGATGAT 3715
Db 3733 AAGGTTCTCAAGAGCCCTGAGCAAGCAAGAGAGCCGACGATGATGATGATGATGATGAT 3792
OY 3716 CACGCTGTGCTACATCCAGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 3775
Db 3793 CACGCTGTGCTACATCCAGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 3852

OY 3776 AAGAGCATGCGACACATCAACACCTGCTGGCCAGAAAGCAATATTTTCCATGCTC 3835
Db 3853 AAGAGCATGCGACACACATCAACACCTGCTGGCCAGAAAGCAATATTTTCCATGCTC 3909
OY 3836 ACTGTCAGGCTGGAGCAAGAGGCTGAGTGAAGTGTGCTGCTGCTGCTGCTGCTGCTGCT 3895
Db 3910 ---GTTGAGGCTGGAGCAAGAGGCTGAGTGAAGTGTGCTGCTGCTGCTGCTGCTGCTGCT 3965
OY 3896 TTAATATTTGTTTAAACATGCAATTAATCAAAAGTTAAAGGTGA 3942
Db 3966 CTAATATTTGTTTAAACATGCAATTAATCAAAAGTTAAAGGTGA 4012

RESULT 12
US-09-880-107-2299
; Sequence 2299, Application US/09880107
; Patent No. US20020142981A1
; GENERAL INFORMATION:
; APPLICANT: Horne, Darci T.
; APPLICANT: Vockley, Joseph G.
; APPLICANT: Scherf, Uwe
; APPLICANT: Gene Logic, Inc.
; TITLE OF INVENTION: Gene Expression Profiles in Liver Cancer
; FILE REFERENCE: 44921-5028-WO
; CURRENT APPLICATION NUMBER: US/09/880,107
; CURRENT FILING DATE: 2001-06-14
; PRIOR APPLICATION NUMBER: US 60/211,379
; PRIOR FILING DATE: 2000-06-14
; PRIOR APPLICATION NUMBER: US 60/237,054
; PRIOR FILING DATE: 2000-10-02
; NUMBER OF SEQ ID NOS: 3950
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 2299
; LENGTH: 3924
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: Genbank Accession No. US20020142981A1 M23234
US-09-880-107-2299

Query Match 52.3%; Score 2240; DB 10; Length 3924;
Best Local Similarity 74.8%; Pred. No. 0;
Matches 2887; Conservative 0; Mismatches 945; Indels 30; Gaps 5;

OY 50 GAGAGAACTTGTGAAATGGCAAAAAAGTAAAAAATGAGAAAGAAAGAAAGAA 109
Db 81 GCGAGAGGAGACTTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAG 140
OY 110 ---CCAAGTGCAGCAGCTTTCGAATGTTTGGCTATTCGAATGTTTGGCTGATAGGTTGAT 166
Db 141 GTGAAATATGATGAGTATTAATGATGATGATGATGATGATGATGATGATGATGATGATGAT 200
OY 167 ATGTTGTTGGGAGCAATGCTGCTGATGATGATGATGATGATGATGATGATGATGATGATGAT 226
Db 201 ATGTTGTTGGGAGCAATGCTGCTGATGATGATGATGATGATGATGATGATGATGATGATGAT 260
OY 227 GTTTTGAAGAAATGACAGATGCTTTCGAATGCAAGAAATTCAGAAAGAAACCAACTTTT 286
Db 261 GTATTTGAGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 314
OY 287 CCAATTAATTAATGAAATATTAAGAAATATTAAGAAATATTAAGAAATATTAAGAAATATTAAG 346
Db 315 CCAATTAATTAATGAAATATTAAGAAATATTAAGAAATATTAAGAAATATTAAGAAATATTAAG 362
OY 347 GAGAAATGACAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 406
Db 363 GAGAAATGACAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 422
OY 407 GCTTACATCAAGTTTCAATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 466
Db 423 GCTTACATCAAGTTTCAATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 482

QY	467	AAACAATTTTTCATGCTATCATGAGACAGAGGAAATGGCTGGTTTGAGCTGATGACGT	526
Db	483	CAGAAATTTTTTTCATGCTATTCATGACAGAGAAATAGGATGGTTTGACATCATAGACACC	542
QY	527	GGGAGGCTTAACACCGGGCTCAGACAGATGTCCTCCAAATCATGAAGAAATGGCGAC	586
Db	543	ACTGACTCATATGCGGGCTACAGATGATGACATCTCCAAATCATGTAAGAAATTTGGTAC	602
QY	587	AAAAATGGAAATGTTCTTTCACCTCAATAGCAACATTTTTCACGGTTTTATATAGGGGTTT	646
Db	603	AAGGTTGGAAATGTTCTTTCAGAGAGTAGCCACGTTTTTTTGAGATTCATATGGGGATTC	662
QY	647	ACAGTGGTTGGAAGCTAACCCCTGTGATTTTGGCCATACGCCCTGTTCTTGACTTTCA	706
Db	663	ATCAGAGAGTAGAAGCTCACCCCTTGATATATGGCCATACGCCCTTATTTAGGACCTCTG	722
QY	707	GCCGCATCTGGGCAAAAGATACATCTTCATTTCTGTATTAAGAACTCTTGGCTATGCA	766
Db	723	GCACCGCTTTGGCCAAAGATCTCTCGCATTTATGTGCAAAAGAACTAGCTGTTATGCA	782
QY	767	AAAGCTGGAGCACTAGCTGAGAAGTCTTTAGCAGCAATCAGAACTGATTTGCTTTGGA	826
Db	783	AAAGCAGGCGCCGTGGCAGAAAGAGCTCTGGGGGCCATCAGAGACTGTGATAGCTTTCCGG	842
QY	827	GGACAAAAGAAAGAACTTGAAGGTACACAACAAAATTTTACAGACCTAAAGAAATGGG	886
Db	843	GGCCACAGACAAAGAGCTGGAAGAGTATCGAAMACATTTTAAAAAATCCCAAAAGATTTGA	902
QY	887	ATAAAGAAAGCTATACGGCCACATTTCTATTTGTCGCCCTTCTTATTTGATCTATGCA	946
Db	903	ATTAAAAAGCTATTTTACGACAAACATTTTCCATGGGATTTGCTCTCTGTTAAATATATCA	962
QY	947	TCATATGCTGTGGCTTTTGTGATTTGAGACCTCCTTGGTCTCTCCAGTAATATTTCTAT	1006
Db	963	TCATATGCACTGGCTTGTGATGATATCCACTGTATGATATCAAAAAGAAATATCTAT	1022
QY	1007	GGACAAGTACTGCTGTCTTCTTCTGATTTAATTTGGGGCTTTAGTATTTGGACAGCA	1066
Db	1023	GGAAATGCATACAGTTTTTTTTTTCATCTTAATTTGAGCTTCAAGTGTGGCCAGGCT	1082
QY	1067	TCCCCAGATTTGAAGCATTTTGGCAACGCAAGAGAGACAGCTTATGAATCTTCAAGATA	1126
Db	1083	GCCCATGATTTGAATGCTTTTGGCCAAATGCAAGAGAGAGCAGCATATGTGATCTTTGATTT	1142
QY	1127	ATTGACATTAACACAGAGATTTGACAGTATTCGACAGTGGACATTAACACAGATPATTT	1186
Db	1143	ATTGATTAATATCCTAAATTTGACAGTTTTTTCAGAGAGGACACAAACCAGACAGCATC	1202
QY	1187	AAGGCAAAATTTGGAATTCAAAATATGTTCACTTCAGTTACCTCTTCGAAAAGAAAGTTAG	1246
Db	1203	AAAGGGAATTTGGAGTTCAATGATGTTTCACTTTCTTACCTCTTCGAGCTAACGTTAAG	1262
QY	1247	ATCTTAAAGGGTCTCACCTCGAAGGTTTCAGAGTGGCAGACAGTGGCGCTGGTTGGAC	1306
Db	1263	ATCTTGAAGGGCTCAACCTGGAAGGGTCAGAGTGGGCGACAGCGTGGCCCTGTTGGAAAT	1322
QY	1307	AGTGGCTGGGGAAGACAGACCGCTGAGCTGATGAGTGCAGAGGGCTATAGACCCACAGAT	1366
Db	1323	AGTGGCTGGGGAAGACACAGCGTCCAGCTGATACAGAGGCTATAGACCCCTGATAG	1382
QY	1367	GGCATGCTCTGATTTGATGGACAGGACATTTAGACCATTAATGTAAGCACTTTGGGAA	1426
Db	1383	GGCAACAATTAACATTTATGAGGCGAGATTTATAGACATTTAATGTAAACTATCGAGGAA	1442
QY	1427	ATTACTGTTGGTGGAGTACGAGCGCTGTGTTTGGCACACAGATGCTGAAAACATT	1486
Db	1443	ATCATTTGTTGTGTGAGTCAAGGCGGGTCTGTTTTCCACACAATTTGCTGAAAATATT	1502
QY	1487	CGCTATGGCCGCAAAATGTCAACATGTGATGATTTGAAAAGCTGTTATAGGAAGCAAT	1546
Db	1503	TGTTATGGCCGTGGAATTAACATGTGATGATTAAGAAAGCTGTCAAGGGGCAAC	1562
QY	1547	GCCATGATTTTATCATGAACCTACTTAATAATTTGACACCTGTGTTGGAGAGAGGG	1606

D	1563	GCCTATGAGCTTTATATCATGAAATTTACACAGAAATTTTACACCCCTGGTTGGAGAGAGG	16
Q	1607	GCCCACTGATGTTGATGAGACAGAAACAGAGATGCCATTGCTGGGCCCTGGTTGCAAC	16
D	1623	GGCCAGCTAGAGTGTGGGCAAGACAGAGATGGCCATTGACAGTCCCTGGTTGCCAAC	16
Q	1667	CCCAAGATTCTTCTGCTGGATGATGAGGCAAGCTCAGCTCTGGACACTGAAAGTAGACAGT	17
D	1683	CCCAAGATTCTTCTGCTGGATGATGAGGCAAGCTCAGCTCTGGACACTGAAAGTAGACAGT	17
Q	1727	GTTTCAGTGGGCCCTGGATTAAGGCCAGAAAGGGCCGACTACCATGTTGATTAAGTATGCT	17
D	1743	GTACAGCAGAGCTGTGGATTAAGGCCAGAAAGGCCGACCACTATGTGATACACACGA	18
Q	1787	TGTGTACAGTTGCTAATGCCGATGTCAATGCTGTGTTTATGATGATGAGATCAATGTGGAG	18
D	1803	CTGTCTAGGCTCCAAATGACAGATGTCATGCTGCGGTTTGAGAGATGAGTAAATGTGGAG	18
Q	1847	AAAGGAATCATGTGATGAACATATGAAAGGAAGAGGCAATTACTTCAACTGTCCAAATG	19
D	1863	CAAGGAACCCACAGCAACTGATGAAAGAAAGGAGGGGTTACTTCAAACTGTCCAAATG	19
Q	1907	CAGCAAGAGGAATGAAATTTGAGTTAGAAATGGCCACTGCTGAATCCAAAGTGAAGT	19
D	1923	CAGCAATCAGGAAGCCAGATCCAGTCCAGTGAAGA-----ATTGGACCTAAATGATGAAAG	19
Q	1967	GATGCTTGGAAATGTCCTCCAAAGATTCAGGGTCCAGTTTAAATPAAAAAGATCAACT	20
D	1977	GCTCCCACTGAATGGCCCCAAATGGCTGGAAATCTGCTAAATTTAGGCAATTCACACAG	20
Q	2027	CGCAGAGTATACATGCACACACAAAGGCCAAGACAAAGACTGTGTAACAAAGAGACTTG	20
D	2037	AAAAACCTTAAAAATTCACAAATGTGTGAGAGACCTGATGTGGAACCGATGACTT	20
Q	2087	AATAGAAATTAACCTCCAGTTTCTCTGAGAGATCTTGAAGCTGAACCTCAACTGAATGG	21
D	2097	GAAACAAATGTGCCACAGCTGTCTTCTTGAAGGTCCTGAAATCGAATTAACACGAAATGG	21
Q	2147	CCTTATTTGTGGTGTATATTTTGTGCTATTAATAACGAGGCGTCGACCAACGAACTT	22
D	2157	CCCTACTTGTGCTGGGAACAGTATGTGCANTGTGCAAATGGGGGGCTTCAGCCGGCAAT	22
Q	2207	TCAATAATATTTTCAAGGATTTAAGGATCTTTACCCGAGATGAGATCTCGAAGCAAAA	22
D	2217	TGAGTATATCTCAGAGATCAATACGATTTTGGACAGGCCATGA--TGCAGTGAAG	22
Q	2267	CGACAGAAAGTACATGTTTCTTATTTGTTTCAGTCCCTGGAAATTAATTTCTTTAT	23
D	2274	CACGAGAAATGCAATATCTCTTTGATTTCTTATTTCTGGGAATTAATTTCTTTTTT	23
Q	2327	ACATTTTCTCCAGGGCTTCATTTTGGCAAAAGCTGGGAGAGATCTCATTAAGCGGCTT	23
D	2334	ACTTTCTCTTCCAGGGTTTACGTTTGGGAAAGCTGGGAGATCTTCACAGAAAGACTG	23
Q	2387	CGATACATGTTTATAGATTCATGTCAGACAGATTCAGCTGTGTTGATGACCTTAA	24
D	2394	CGGTCAATGGCTTTAAAGCAATGCTTAAGACAGGACATGAGCTGTGTTATGACATAA	24
Q	2447	AACACCATGAGCACTTGACAAACAGGCTTGCCAAATGATCGGCTCAAGTTAAAGGGCT	25
D	2454	AACAGTACTGTGCACTTCTACAAAGACTTGCACAGATCTGCCAATCCAGAGAGCC	25
Q	2507	ATAGGTTCCAGGCTTGCCTATTAACCAAAATATATAGCAATTTTGGACAGGCAATAT	25
D	2514	ACAGGAACCCAGGTGGCTTTATATGCAACAAATATATGCTACCACTGTATTTTC	25
Q	2567	ATATCTTAATCTATGTTGGCAATTAACACTTTTACTCTTACCAATTTTACCACTAT	26
D	2574	ATATCTTAATCTATGTTGGCAATTAACACTTTTACTCTTACCAATTTTACCACTAT	26
Q	2627	GCATATGAGAGATTTTGAATGAAATGTGTCTGGAACAGCACTGAAGATTAAGAA	26

Db 2634 GCTGTGTGAGGAAATTTGTAATGAATGTGTGGTGAATGCCAAAAGAGATAAAAA 2693
Qy 2687 GAGCTAGAGAGAGCTGGGAGATGTCTACAGAGCAATGCCAACTGTGT 2746
Db 2694 GAACGTGAGAGTGTGGAAGATTTGCAACAGAGCAATGAAAAATATTAAGAGTTGG 2753
Qy 2747 TCTTTGACCTGGAGAGAAATTTGAATACATATGACACAGAGTTTGCATAGCATTAC 2806
Db 2754 TCTTTGACCTGGAGAGAAATTTGAATACATATGAAAAATTTGATGAGACTTAC 2813
Qy 2807 AGAAACTTTTGAAGAAAGACACATCTTGGGGTCTCATTTTCTATCCAGGCAATG 2866
Db 2814 AGGAATTTCTGACAGAAAGACACATCTATGGAATTAATTTAGTATCTACAGACTTT 2873
Qy 2867 ATGTATTTTCTATGCTGGCTTTCGGTTTCCGTTTGTGCTCTACTTGTGCGCAATGTTCC 2926
Db 2874 ATGTATTTTCTATGCTGGCTTTCGGTTTCCGTTTGTGCTCTACTTGTGCGCAATGTTCC 2933
Qy 2927 ATGAACCTTTCAAGATTTCTTTTGTATTTCTACAGCTATTTGTTTGTGCTGCAATG 2986
Db 2934 ATGGCTTCAAGATTTCTTTTGTATTTCTGCAATTTGATTTGGTGAAGTGGCTCTA 2993
Qy 2987 GGGCAGGTAGTTCAATTTGCTTCCATATGCAAGCAAGCAAGTATGACAGCCACGTC 3046
Db 2994 GGCACATGCCAGTTCAATTTGCTTCCATATGCTAAGCTAAGCTGTGTCGACGCCACTTA 3053
Qy 3047 ATCATGATCATTTGAAAAAGCCCTCTGATTTGACAGCTTACAGCCCTCAGCCCTCAAGCA 3106
Db 3054 TTTATCTCTTTTGAAGAAAGCAACCTCTGATTTGACAGCTTACAGAGGGGCTCAAGCTT 3113
Qy 3107 AATACCTTTGAAGAAAGTATGACATTTATATGAGTGTGCTTCACTATCCACTGACCA 3166
Db 3114 GATTAATTTTGAAGAAATTTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 3173
Qy 3167 GATATCCCGTGTCCAGGGCTGAGCCCTGAGGTGAAGAGGCGCAGACCTGTGCTC 3226
Db 3174 AAGCTCCAGTGTTCAGGGGCTGAGCCCTGAGGTGAAGAGGCGCAGACCTGTGCTC 3233
Qy 3227 GTAGTAGCAGTGTGTGGGAAAGACACAGTTTTCAGCTTCTTACAGCGCTTATGAC 3286
Db 3234 GTGGGCGAGTGTGTGGGAAAGACACAGTGTGTCCAGCTCTGAGAGCGTTTCTACAC 3293
Qy 3287 CCCGTGCTGGTTCAGTGTATTTGATGCAAGATTAAGATTAAGATTAAGATTAAGATTA 3346
Db 3294 CCCGTGCTGGTTCAGTGTATTTGATGCAAGATTAAGATTAAGATTAAGATTAAGATTA 3353
Qy 3347 CTCGAGACACACCTGGGATGTGTCTCAGAGAGCCATCTCTTGTGACGACAGTTGCC 3406
Db 3354 CTCAGAGCTCAACTCGGAATCGTGTCTCAGAGAGCCATCTCTTGTGACGACAGTTGCC 3413
Qy 3407 GAGAACTTGTCTATGAGAGACAGCGGGTGTATCAATGAGATTAATGAGATTAATGAGCA 3466
Db 3414 GAGAACTTGTCTATGAGAGACAGCGGGTGTATCAATGAGATTAATGAGATTAATGAGCA 3473
Qy 3467 GCCAAGAGAGGCAACATACACACTTCATGAGACACTCCCTGAGAAATTAACACACCA 3526
Db 3474 GCCAAGAGGCAACATACATCTTTCATGAGAGCTTTACCCCAAAATTAAGAAACAA 3533
Qy 3527 GTAGAGACAAAGAACCCAGCTCTGAGAGGTTCAAAAACAGAGANTTGTATTTGCCCA 3586
Db 3534 GTGGAGATTAAGAGAGCTCAGCTCTCAGAGGTTCAAAAACAGAGANTTGTATTTGCCCA 3593
Qy 3587 GCTCTTTTGAAGAGCTCATATTTTCTTTTGTATGATGATGATGATGATGATGATGATGAT 3646
Db 3594 GCTCTTTTGAAGAGCTCATATTTTCTTTTGTATGATGATGATGATGATGATGATGATGAT 3653
Qy 3647 GAAAGTGAAGAGTTGTCCAGAGAGCCCTGAGCAAAAGCCGAGAGAGGCGCACCTGAT 3706
Db 3654 GAAAGTGAAGAGTTGTGTCCAGAGAGCCCTGAGCAAAAGCCGAGAGAGGCGCACCTGAT 3713
Qy 3707 GTGATGCGCCAGCGCTTGTGTCCAGAGAGCCCTGAGCAAAAGCCGAGAGAGGCGCACCTGAT 3766
Db 3714 GTGATGCGCTACCGCTTGTGTCCAGAGAGCCCTGAGCAAAAGCCGAGAGAGGCGCACCTGAT 3773

Qy 3767 GCGAAGTCAAGAGGATGCGACATCAACAGCTCTGCGCCAGAAAGCATCTATTTT 3826
Db 3774 GCGAGAGTCAAGAGGATGCGACATCAACAGCTCTGCGCCAGAAAGCATCTATTTT 3833
Qy 3827 TCCATGCTCAGTGTCCAGGCTGAGCAAAAGCCCTGATGAACTGTGCGCCATGAGCTGTT 3886
Db 3834 TCAATGCTCAGTGTCCAGGCTGAGCAAAAGCCCTGATGAACTGTGCGCCATGAGCTGTT 3893
Qy 3887 AAATATTTTAAATATTTGT 3908
Db 3894 AAAATTAATTAATATTTCT 3915

RESULT 13

US-09-917-800A-1560
Sequence 1560, Application US/09917800A
Patent No. US20020119462A1
GENERAL INFORMATION:
APPLICANT: Mendrick, Donna
APPLICANT: Porter, Mark
APPLICANT: Johnson, Kory
APPLICANT: Castle, Arthur
APPLICANT: Elshoff, Michael
TITLE OF INVENTION: Molecular Toxicology Modeling
FILE REFERENCE: 44921-5038-US
CURRENT APPLICATION NUMBER: US/09/917,800A
PRIORITY FILING DATE: 2001-07-31
PRIORITY APPLICATION NUMBER: US 60/222,040
PRIORITY FILING DATE: 2000-07-31
PRIORITY APPLICATION NUMBER: US 60/222,880
PRIORITY FILING DATE: 2000-11-02
PRIORITY APPLICATION NUMBER: US 60/290,029
PRIORITY FILING DATE: 2001-05-11
PRIORITY APPLICATION NUMBER: US 60/290,645
PRIORITY FILING DATE: 2001-05-15
PRIORITY APPLICATION NUMBER: US 60/292,336
PRIORITY FILING DATE: 2001-05-22
PRIORITY APPLICATION NUMBER: US 60/295,798
PRIORITY FILING DATE: 2001-06-06
PRIORITY APPLICATION NUMBER: US 60/297,457
PRIORITY FILING DATE: 2001-06-13
PRIORITY APPLICATION NUMBER: US 60/298,884
PRIORITY FILING DATE: 2001-06-19
PRIORITY APPLICATION NUMBER: US 60/303,459
NUMBER OF SEQ ID NOS: 1740
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 1560
LENGTH: 3912
TYPE: DNA
ORGANISM: Rattus norvegicus
FEATURE:
OTHER INFORMATION: Genbank Accession No. US20020119462A1 NM_012690
US-09-917-800A-1560

Query Match 48.7%; Score 2081.8; DB 10; Length 3912;
Best Local Similarity 71.6%; Pred. No. 0;
Matches 2798; Conservative 0; Mismatches 1087; Indels 24; Gaps 4;

Qy 3 AGCGAGAGTGTGGAGATGCTGTGAGAGAGCCGTAAGGAGCTGACAGAGAAAGTCT 62
Db 25 AGCGTGAAGATGATCTGTGAGAGAGCAAGAAACGGAACAGCGCGCCCTGACGCGACT 84
Qy 63 GGAAGATGGCAAAAAAGTAAAAAATGAGAAAGAAAGAAACCAACTGTCA--- 119
Db 85 TTGAAGTGAAGAGATGAGCAACAGAGAGAGAAAAAGAAAGTAAATTAATG 144
Qy 120 GCAAGTGTCAATGTGTGCTATTTCAATTTGCTTGTATGAGTTGTATGTTGGTGGCA 179
Db 145 GCGCTTACACCTGTGTGATCTGTGATTTGGCAGATTAATTTGTTATGCTCTGGCA 204

QY	180	CAATGGCTGCAATCATCCATGAGCTGCACCTCCCTCATGATGCTGGTTTGGAAACA	239
Db	205	CGCCCATGGCGATAGCTACAGGATCAGGCTTCCCTTATGATGATCTTTGGAGAA	264
QY	240	TGACAGATAGCTTGGCAAAATGAGAAATTCAGAAAGAAACAACTTTCCAGTTATTA	299
Db	265	TGACATATAGTTTGTAGT-----AATGCGGGAACCTTTTCCCTGGCAGT-----	310
QY	300	ATGAAGATTATAGCAACATATACAAATTTTCATCAACATCTCGAGAGGAATGACA	359
Db	311	----GATTTTTCATTTGTCATATGTAATTCAGAAAGAAATCTCGAAGAAATATGACTA	366
QY	360	CGTATGCCATATATATACAGTGGGATCGGTGCTGGCGTCTGGTCTTACATCCAG	419
Db	367	GATATGACTACTACTATTCGGAGCTAGGTGGATGTTCTTTTGGCTGTATATCCAG	426
QY	420	TTTTCATTCGTGCTCGGCACAGAGAAAGCATATCAAAATTAAGAAACATTTTTC	479
Db	427	TCTCCTCTGGACTTTGGCAGCTGGCCGACAAATAGAGAAATCAGCAAAATTTTTC	486
QY	480	ATGCTATATGCGAGAGAGATTTGGCTGTTTGAACGTGATGAGTGGGAGCTTAACA	539
Db	487	AGCCATCTCTTGACAAAGAAATGGGCTGTTGATATCAAGGGAGACACGACTCAACA	546
QY	540	CCCGGCTACAGAGATGCTCCAAATATGAAGAAATGGCGACAAATTTGGAATG	599
Db	547	CGCGGCTACAGATGACATCTCCAAATATGAGAAATTTGGTGAACAGTTTGAATG	606
QY	600	TCTTTCATCTAATGACAACTTTTTCACCGTTTATAGTGGGTTTACACGTGTGA	659
Db	607	TCTTTCAGCAAGATAGCCAGTTTTTCAGAGATTCATAGTGGGTTTACATAGAGCTGA	666
QY	660	AGCTAACCCCTTGATTTTGGCCATCAGCCCTTCTTGAGCTTTCAGCCCATCTGG	719
Db	667	AATCAGCCCTGATATGATGCGCATACCGCATCTTGAGGCTCTCTACAGCCGTTGG	726
QY	720	CAAGATACTATTTTACTTACTTAAGAACTCTGGCCTATGAGAAAGCTGAGCAG	779
Db	727	CAAGATACTCTCAACATTCAGTGAAGAAAGTACTGCTCTATGCAAAAGAGAGTGGC	786
QY	780	TACTGAAAGATCTTAGCAGCAATCAGAACTGTGTATCCCTTTGGAGACAAAGAAAG	839
Db	787	TGCGGAAAGAGGTCTGGAGAGCCATCAGAGCCGTGATAGCTTTGGGGGCCAGAAAG	846
QY	840	AACCTAAAGGTGCACAAAAATTTAGAGAAAGCTTAAGAAATTTGGATTAAGAAAGCTA	899
Db	847	AGCTAAAGGTATCGAAGAGCAATTAAGAAATGCAAAAGAGTGAATTAAGAAAGCTA	906
QY	900	TACGGCCAACTTTCTATTGTGCGCTTCTTATGATCTATGCAATCATCTGCG	959
Db	907	TCTCGGCCAACTCTCATGGGCATTTGCTTTTGTATATATGACATCTTATCCACTG	966
QY	960	CTTTCGATGAGGAGCTCTGTCCTCTCAGTGAATTTCTATTGGACAACTACTCA	1019
Db	967	CTTTCGATGAGATCTCCTGTGTTATATCAAAAGAAATATACAAATTTGAAATCCATGA	1026
QY	1020	CTGCTCTCTTCTCTATTAATTTGGGCTTTTATGATTTGGACAGGATCCCCAAGCTG	1079
Db	1027	CAGTGTCTTCTCATCTCATTTGGGCTTCACTGTGGGCAAGGCGCCCTGATTTG	1086
QY	1080	AAGCATTTGGCAAGCAAGAGAGCAGCTTATGAATCTTCAGATATATGACAAATTAAC	1139
Db	1087	ATGCTTCCCAATGCTAGAGAGCAGCCTATGTATCTTGAACATATATGATATATATC	1146
QY	1140	CAAGCATGAGCATATTCGAGAGTGAATTAACCAAGATATATTAAGGAAATTTGG	1199
Db	1147	CTAAATTTGACAGTTTTCAGAGAGAGACAAAGCCAGACACATCAAAAGAAATTTGG	1206
QY	1200	AATTCAAAATGTTCACTCAGTTACCTCTGAAAAAGAGTTAAGATCTTAAGAGGTC	1259
Db	1207	AGTTCAAGTACGTTCTTCTTCCATCCCATCTGGGCTAATATCAAGATCTTGAAGGSC	1266
QY	1260	TCAACCTGAAGTTCAAGTGGGCGAGCAGTGGCGCTGTTGGGAAACAGTGGCTGGGGA	1319

Db	1267	TCAACCTGAAGGTGAAGAGCGGGCAGACGGTATGCCCTGTTGGCAACAGTGGCTGGGA	1330
QY	1320	AGAGCAGACCGTGCAGCGATGACAGAGCTCTATATGCCACAGATGGATGTCTGTA	1337
Db	1327	AAAGCACAACCTGTCCAGCTGTGCGAGAGGCTCTTACGACCCACAGAGGATAGATTACA	1380
QY	1380	TTGATGAGAGACATTTAGACCAATAAATGTAAAGCATCTTGGGAATTTCTGTGTGG	1447
Db	1387	TGATGGGACAGACATCCGGAACTTTAAGCTAGGTCTTAAGGAATTCATCGCGGTGG	1454
QY	1440	TGAGTCAGAGAGCTGTGTGTTTGGCCACAGATAGCTGAAGAAATTCGTATGGCGCG	1497
Db	1447	TGAGTCAAAGCGGCTAGCTTCTTACACAGATTTCTGAAGAAATATCCGATATGGCGTG	1504
QY	1500	AAAAATGACCATGATGAGATTGGAAGAGCTTTAAGAAAGCCAAATGCTATGATTTTA	1557
Db	1507	GGAATGTAAACATGATATAGATTAAAGAAAGCTGTCAAGAGAGCTTAAGCTATGACTTCA	1564
QY	1560	TCATTAACCTACCTAATTAATTTGACACTCTGTTTGGAGAGAGAGGGGCCAGCTGATG	1617
Db	1567	TCATTAACCTGACAGAAATTTGACACCTGTTGTGTGAGACAGAGGGGCCAGCTGATG	1624
QY	1620	GTGGACAGAAACAGAGATGGCCATTTGCTGGGCGCTGTGTCGCAACCCAGATTTCTTC	1677
Db	1627	GGGGACAGAAACAGAGATGGCCATTTGCTGTGCTGTGCTGTGTCGCAACCCAGATTTCTTC	1684
QY	1680	TGCTGGATGAGGCAACGCTCAGCTGTGACACTGTAAGTGAAGAGAGTGTTCAGTGGGCC	1737
Db	1687	TGCTGGACAGAGCCACGTCACACCTTTGGACACAAAGACGAGTGTGAGCAGCGCTGC	1744
QY	1740	TGGATTAAGGCGCAAAAAGGCCCGGACTACCATTTGTATAGCTCATCTGTTGTACAGTTTC	1797
Db	1747	TGGATTAAGGCGCAAAAAGGCCCGGACACCATTTGTATAGCTCATCTGTTGTACAGTTTC	1804
QY	1800	GTAATGCCAGATTCATTTGCTGTGTTTGAATGATGATGATGATGATGATGATGATGATG	1857
Db	1807	GGAATGCAATGTCATCTGCTGCTGTTTGAAGATGCTCATCTGTTGAGCAAGAGACCTACA	1864
QY	1860	ATGAACTCATGAAGAGAGAGAGGCAATTTACTTCAAAATTTCAAAATGAGAGAGAGAGAA	1917
Db	1867	GTAAGCTGATTAAGAGAGAGAGGATCTACTTCAAGATTTTAAACATGCAAGCATAGAGAA	1924
QY	1920	ATGAATATGATTAAGAAATGGCACCTGTGTAATCCAAAGTGAAGATGATGCTTGGAAA	1977
Db	1927	GCCAGATCTCTGAGAGAAATTTGAAGTGTGACTTAAGTGAATGAAGAGCTGCTGGAGTG	1984
QY	1980	TGCTCCAAAGATTCAGAGCTCAGTTTAAATTAAGAAAGATCAACACTGCGAGAGTATAC	2039
Db	1987	TGGCCCCAATATGAGTGAAGAGCAGCATTTTAAGAAATTTCAAGAGAGAGAGTCTGAAA	2046
QY	2040	ATGCACACAAAGCCACAGACAGAAAGCTTGTGAACAAAGAGAGCTTGAATGAGATGATAC	2099
Db	2047	GTTACAGGGCGCATTAATATAGCTGATGTGAGAAACCAATGAACATGATGATCAACCTGC	2106
QY	2100	CTCCAGTTTCTCTTGGAGAGATTTGAGAGCTGAGACTCAACTGAATGGCTTATTTTGG	2159
Db	2107	C	

Db 2344 AGGTTTCACATTCCTGGGAAAGCTGGGAGATCTCTACCAAGGCTCCGGTCCATGGCT 2403
OY 2400 TCAGATCCATGTCAGACAGAGATGTAGCTGGTTGATGATGACCTTAAACACACTGGAG 2459
Db 2404 TCAGAGCAATGCTAAGACAGAGATAGCTGGTTGATGATGATGATGATGATGATGATG 2463
OY 2460 CATTGACACAGGCTTGGCAATGATGCGGCTCACTTAAAGGGCTATGAGTCCAGGC 2519
Db 2464 CCTCTCTACAGAGCTGCGACAGAGCTGCGAGGCTCCAGAGGCGCACAGAGACAGT 2523
OY 2520 TTGGTGCATTAACCAATATATGCAATCTTGGGACAGGCAATTAATATATCTTAATCT 2579
Db 2524 TGGCTTTAATGTCACAGACAGACAGGCACTTGGAAAGGGTATTAATATATTAATTT 2583
OY 2580 ATGGTTGGCAATTAACACTTTTACTCTTACCAATTTGACCAATTCATTCGAAATGCGAG 2639
Db 2584 ACGGTTGGCACTGACACTTCTGCTTATCAGTTGTTTCATTCAATGCTGTAGCGGGA 2643
OY 2640 TTGTTGAATGAATAATGTTGCTGACAGACACTGAAGAGTAAGAAAGAGATGGAAGAG 2699
Db 2644 TTGTTGAATGAATAATGTTGCTGACAGACACTGAAGAGTAAGAAAGAGATGGAAGAG 2703
OY 2700 CTGGAGATTTGCTACAGAGACCATCGAAAATCTCCGACTGTTGTTCTTGTACTCGGG 2759
Db 2704 CTGGAGATTTGCTACAGAGACCATCGAAAATCTCCGACTGTTGTTCTTGTACTCGGG 2763
OY 2760 ACAGAAATTTGAATACATGATATGACAGAGATTTGCAAGTACATACAGAACTCTTTGA 2819
Db 2764 ACAGAAATTTGAATACATGATATGACAGAGATTTGCAAGTACATACAGAACTCTTTGA 2823
OY 2820 GGAAGACACACATCTCTGGGGCTCTCATTTTCTATACACAGGCAATGATATTTCT 2879
Db 2824 GGAAGCTCACATCTGACATCACTTTAGATCTCAAGCATCTGACTTCTT 2883
OY 2880 ATGCTGGCTGTTCCGGTTTGGTGGCTTGGTGGCAATGAGTCAATCTTCAAG 2939
Db 2884 ATGCTGGCTGTTCCGGTTTGGTGGCTTGGTGGCAATGAGTCAATCTTCAAG 2943
OY 2940 ATGTTCTTTGATATCTGATGATGATGATGATGATGATGATGATGATGATGATGATG 2999
Db 2944 ATGTTCTTTGATATCTGATGATGATGATGATGATGATGATGATGATGATGATGATG 3003
OY 3000 CATTGCTCTGACTATGCAAAAGCAAGTATGACAGGCACTGATCATGATCTG 3059
Db 3004 CATTGCTCTGACTATGCAAAAGCAAGTATGACAGGCACTGATCATGATCTG 3063
OY 3060 AAAAAACCTCTGATTTGACAGCTACAGCCCTCAGGGCTCAAGCCCAATATCTGGAAG 3119
Db 3064 AAAAAACCTCTGATTTGACAGCTACAGCCCTCAGGGCTCAAGCCCAATATCTGGAAG 3123
OY 3120 GAAATGTCATTTAATGAGTGTGTTCACTATCCCACTGACACAGATCCCGCTGC 3179
Db 3124 GAACTGTGACATTTAATGAGTGTGTTCACTATCCCACTGACACAGATCCCGCTGC 3183
OY 3180 TCCAGGGCTGAGGCTGAGGTGAGAGAGGCGCAGAGCTGGCCCTGATAGTACAGTG 3239
Db 3184 TCCAGGGCTGAGGCTGAGGTGAGAGAGGCGCAGAGCTGGCCCTGATAGTACAGTG 3243
OY 3240 GCTGTGGAGAGACAGAGTGTCTAGCTCTAGAGGGCTCTTGAAGCCCTTGGAGTGT 3299
Db 3244 GCTGTGGAGAGAGACAGAGTGTCTAGCTCTAGAGGGCTCTTGAAGCCCTTGGAGTGT 3303
OY 3300 CAGTGTCTATGATGGGCAAGATTAAGACACTGATGTCTCAAGGCTCCAGACACC 3359
Db 3304 CAGTGTCTCTGATGTGTAGAGAAACAAATGTCATGTCAGTGTCCGAGCTCAAC 3363
OY 3360 TGGGCAATGCTCTCAGAGAGCCATCTGTTTGAATCAGACATTTGCCGAAACATTTGCT 3419
Db 3364 TGGGCAATGCTCTCAGAGAGCCATCTGTTTGAATCAGACATTTGCCGAAACATTTGCT 3423
OY 3420 ATGAGAGAAACAGCGGGTGTATCAGATGAAGATTTATGAGAGGCAAGAGGCCA 3479
Db 3424 ACGGAGAAACAGCGGGTGTATCAGATGAAGATTTATGAGAGGCGCCAAAGAGGCCA 3483

OY 3480 ACATACAGCACTTCATCGACACTCTCCAGAAATTAACACACAGATAGACAAAG 3539
Db 3484 ACATACAGCACTTCATCGACACTCTCCAGAAATTAACACACAGATAGACAAAG 3543
OY 3540 GAACCCAGCTCTCTGAGGAGCAAGACAGGATTTGCCATTCAGCTCTTGTAGAC 3599
Db 3544 GAGACAGCTCTCTGAGGAGCAAGACAGGATTTGCCATTCAGCTCTTGTAGAC 3603
OY 3600 AGCTCATATTTTGTCTTTGATGAAGTACATCAGCTCTGATACAGAAAGTGAAGAG 3659
Db 3604 AGCTCTGGGCTCTGATGATGAGCCAGCAGCTCTGATGAGCACTGAGAGTGAAGAG 3663
OY 3660 TTGTCAGAAAGCCCTGGAGAAAGCCAGAGAGCCGACCTGATGATGATGCCAAC 3719
Db 3664 TTGTCAGAAAGCCCTGGAGAAAGCCAGAGAGCCGACCTGATGATGATGCCAAC 3723
OY 3720 GCTTGTCCACCATCCAGAAATGAGATTTAATGATGTTTCAAGATGGCAAGCAAG 3779
Db 3724 GCTTGTCCACCATCCAGAAATGAGATTTAATGATGTTTCAAGATGGCAAGCAAG 3783
OY 3780 AGCATGCAATCAACAGAGTGTGCGCCAGAAAGCATCTATTTTCCATGCTCAGTG 3839
Db 3784 AGCAGGAGCCACACAGAGTGTGCGCCAGAAAGCATCTATTTTCCATGCTCAGTG 3843
OY 3840 TCCAGGCTGAGAAAGCGTATGAACTGTGCCATATGACCTGTTAATATTTTAA 3899
Db 3844 TCCAGGCTGAGAAAGCGTATGAACTGTGCCATATGACCTGTTAATATTTTAA 3903
OY 3900 TATTTGCT 3908
Db 3904 TATTTGCT 3912

RESULT 14
US-09-873-409-14
; Sequence 14, Application US/09873409
; Patent No. US20020037522A1
; GENERAL INFORMATION:
; APPLICANT: Frank, Markus
; APPLICANT: Sayegh, Mohamed
; TITLE OF INVENTION: A Gene Encoding a Multidrug Resistance Human P-Glycoprotein
; FILE REFERENCE: 81994/268611
; CURRENT APPLICATION NUMBER: US/09/873, 409
; CURRENT FILING DATE: 2001-06-05
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 14
; LENGTH: 3621
; TYPE: DNA
; ORGANISM: Homo sapiens
; US-09-873-409-14

Query Match 30.9%; Score 1323.8; DB 10; Length 3621;
Best Local Similarity 62.6%; Pred. No. 2.8e-293; Indels 51; Gaps 6;
Matches 2189; Conservative 0; Mismatches 1257;

OY 418 GGTTCATTTCTGCTGCTGACAGAGACAGATATCAAAATTAAGAAACATTTT 477
Db 162 CAGGTTCATTTCTGCTGCTGACAGAGACAGATATCAAAATTAAGAAACATTTT 477
OY 222 GATTTCTCTTGATTTAATCTGACAGCAGACAGACAGAGATTTGAAACAGTTT 281
Db 222 GATTTCTCTTGATTTAATCTGACAGCAGACAGACAGAGATTTGAAACAGTTT 281
OY 478 TCATGCTATCATGAGAGAGATTTGCTGATGATGATGATGATGATGATGATGATG 537
Db 282 TCATGCTATTTTGGACAGAGAGATTTGCTGATGATGATGATGATGATGATGATG 537
OY 538 CACCGGCTCAGACAGATGCTCCAAATCAATGAAGATTTGGAGCAAAATTTGAAT 597
Db 342 CACTGCTGATGACAGATTTGCTCCAAATCAATGAAGATTTGGAGCAAAATTTGAAT 597

QY	598	GTTCTTTCAC	TCAATAGCAACATTTTTCACCGGGTTTATAGTGGGTTTACACTGGTTG	657
Db	399	GTGTGTCAAAA	CAGTCTACTTTTGATTTGGCCCTGGCAGATGGTTTGGTGAAGGGCTG	458
QY	658	GAAGCTAAC	CCCTGTGATTTTGGCCATCAGCCCTGTGTTGGACATTCACGCCGACATCTG	717
Db	459	GAACTCAC	CCCTACTGTATTCACAGTCTCTCTTAAATAGGCTTCAGGGGAGCATG	518
QY	718	GGCAAGAT	TACTATCTTATCTTACTATTAAGACATCTGGCCATAGCAAACTGGAG	777
Db	519	TTCTAGAT	GTGTCATTCATGTACAGATAGCAATTAAGTGCCTATTCGCAAACTGGGCG	578
QY	778	AGTAGCTGA	AAACTCTTACACCAATCAGAACTGATTTGCCCTTGGAGAGCAAAAGAA	837
Db	579	TGTGGCA	GAAGAACTCTTGTCATCATCCGAAAGTCATACCTTTTGGGCCGAGAGAA	638
QY	838	AGAACTTGA	AAAGGTACAAACAAATTTTAAAGAACTTAAGAAATTTGGCTAAAGAAAGC	897
Db	639	AGAACTTCA	AAAGGTATACAGACAGATCTCAAGAGATCCAAAGATTTTGGCTAAAGAAAGC	698
QY	898	TATCAGG	CCCAATTTCTTATTTGGTCCGCTTTCTTATATGATCATGACATATCTCT	957
Db	699	TATAGCTT	CAAAAGTCTCTTGGTGTGTACTCTTTATGATAGAAACCTATGACT	758
QY	958	GGCTTTC	GTATGAGGACCTCTGTGGCTCTCCAG-----TGAATATTCTATTGGACA	1011
Db	759	TGCTTTT	GTATGAGACCTCTTATTTCTTAAATGGAACCTGGATATATCATCGGAGC	818
QY	1012	AGTATC	CACGTCTCTTCTTCTTCTTATTAATTTAGGGGCTTTTATGATTTGACAGCATATCCG	1071
Db	819	TGTCTT	GTCTTCTTCTTCTTCTTCTTATAGTATATCATAGACAGTTATGATTTGGAGCAGCTCC	878
QY	1072	AAGCATTT	GAAGATTTGCAAAAGCAAGAGAGAGACGTTATGAAATCTTCAGATTAATTTG	1131
Db	879	TCACTTT	GAACCTTCGCAATATGCCGAGAGCTGCTTCTTATATTTTCCAGGTTATTTGA	938
QY	1132	CATTAAC	CACACATGTACAGCTATTTGGAAGAGTGCATTAACAGATTAATTAATTAAGG	1191
Db	939	TAAGAA	ACCAAGTATAGATTAATCTTTTCCAGCTGATTAACCTGAAATCCATATGAAAG	998
QY	1192	AAATTTG	GAATTCAAAAATGATCTTCACTCATATTCCTTCTGAAAGAAATTAAGATCTT	1251
Db	999	AAGTGG	AATTTAAATATGTTTCTTCAATTTTCCATCAACACCATCTATACAGATTTCT	1058
QY	1252	AAAGGCT	CAACCTGAAAGTTGACAGATGAGATGAGGACAGATGGGGCTGTGGGAACACTGG	1311
Db	1059	GAAAGCT	GTGAATCTCAGAAATTAAGTCTGAGAGACAGTCCGCTTGGCTCTCAATGG	1118
QY	1312	CTGCGG	GAAGCAAGCAGCTGTGACGTATGACAGAGGCTCTATGACCCACAGATGGCAT	1371
Db	1119	CAGTGG	GAAGATGCGTACTCCAGCTTTGTGAGAGGTTATATATCCGATGATGGCTT	1178
QY	1372	GGTCTG	ATTTGATGACAGCAATTTGAGCACTTAATTTAAGGACATCTTGGGAAATTTAC	1431
Db	1179	TATCAT	TGGTATGATGAGAAATACATCAGACCTTAAATGTGGCGCATATATGAGAACCATAT	1238
QY	1432	TGTGTGT	GTGATGAGAGAGGCTGTGTGTTTGCACACAGATATGCTTAAGAACTTGTGCTTA	1491
Db	1239	TGAGAT	GTGATAGTCAAGAGCTGTGTTTTTGTGGGACACATCAGTAACTATACAGTA	1298
QY	1492	TGGCCG	CGAAATGTCAACATGATGAGATTGAGAAAGCTGTTTAAGGAAGCCATGGCTTA	1551
Db	1299	TGGACG	AGATGATGATGATGAAAGATGGAAGATGGAAGACACAAAGGAAAGCAAAATGGCTTA	1358
QY	1552	TGATTTA	TATCATCTACCTTAATTAATTTTGGACATCTGTGTTGGAGAGAGAGGGGCCCA	1611
Db	1359	TGATTTT	TATCATCTAGGATTTCTCAATTAATTTAATACATTTGTTAGGGGAAAAAGGAGCTCA	1418
QY	1612	GCTGAG	TGTGTGAGAAACAGAGATGTGCATGTGCTGGGGCCCTGTGTCGACCCCAA	1671
Db	1419	AATGAG	TGAGAGGCAAGAGATTCGATTTGCTGTGGCTTATGTTTGAAGCCCAA	1478

QY	1672	GATTCTTCTGCTGGATGAGGCAACGTCAGCTCTGACACTGAAAGAGAAAGCATGGTTC	1731
Db	1479	GATTCTGATTTTATGATGAGGCTACGTCGTGCCCTGGATTTCAGAAACCAATGCTACGCTGTCA	1538
QY	1732	GGTGGCCCTGGATTAAGGCCGCAAAAGGCCGACATACATTTGTGATAGCTCATGTTGTCTC	1791
Db	1539	AGCTGCACCTGGAAAGGCGAGCAAAAGGTGGGACTCAAAATCGGTGTAGCACACCCGACTTTC	1598
QY	1792	TACAGTTCTGTAATGCCGATGTCATTTCTGTTGTGATGTTGAGATCATTTGGGAGAAAG	1851
Db	1599	TACTATTCCAAAGTCGAGATTTGATTTGTGCACCTTAAGATGTGAATGCTGGCGGAGAAAG	1658
QY	1852	AAATTCATGATGATACCTCATGATAAAGAGAGGCGCATTTACTTCAAACTGTGCACATGCGAC	1911
Db	1659	AGCACATGCTGATCATATGCGCAAAACGAGGTCTATATTTATTCCTTGTGTGATGTCACAG	1718
QY	1912	MAGAGAAATGGAATTTGATTTAGAAATTAAGCCCTGCTGTAATCCAAATGGAAGTGTGC	1971
Db	1719	TATTAATAAAGCGTATGTA-ACAGATGAGCTCAATGACATATTCTACTGAAAGAAAGCA	1777
QY	1972	CTTGGAAATGCTCCAAAAGATTTACGGGTCCAGTTTATATTAATAAAGAGATCAACTGCGAC	2031
Db	1778	ACTCACTCTCTGCACTCTGTGAAG------CATCAAGTCAGACTTCATTTAGCA	1829
QY	2032	GAGTATACATGCAACCAAGGCGAAAGACAGAAAGCTGTGTACAAAGAGACCTTGAATGA	2091
Db	1830	GGCTGAGGAATCCACCAGATCTMAAGATTAAGTCTT-----	1866
QY	2092	GAAATGACTCCAGTTTCTCTGAGAGATTTGMAAGCTGMACTGACCTCAATGATAGGCTTTA	2151
Db	1867	-----CCTGAAGTCTCTATTATTAATTAATTAAGTTTAACAGACCTGAAAGGCTTT	1919
QY	2152	TTTGTGTTGGTATTTATTTTGTCTATTATTAACGGAGGCTCCACACGACATTTTCAAT	2211
Db	1920	TGTGTTCTGGGACATTTGGTCTCTGTCTTAATGTGAACCTGTTCATCCAGTATTTTCCAT	1979
QY	2212	AATATTTTCAAGCATTTATAGGATCTTTACCGAGATGTAGAGATCCGGAAGCAAAACGACA	2271
Db	1980	CATCTTTGCAAAATTTATTAACCATGTTT---GGAATTAATGATTAACCATTTTAAGCA	2036
QY	2272	GAAATGATACATGTTTCTGATTTGTTTTCAGTCCCTGGAATTTATTTCTTTTATTAACAT	2331
Db	2037	TGATGCAGAATATTATTCATGATATTCTGCTATTTTGGGTGTTATTTGGTGTGCTGTTA	2096
QY	2332	TTTTCTCCAGGCTCTACATTTTGGCAAAAGCTGGGAGATCCCTCACTAAAGGCGCTCGTGA	2391
Db	2097	TTTTCATGCAAGGATTTATTTACGGCGAGACAGAGGGAATTTTAAGATGAGATTTAAGCA	2156
QY	2392	CATGTTTTCAGATTCATGCTGTGAGACAGATGTCAAGCTGTGTTGATGACCTTAAAAAC	2451
Db	2157	CTTGGCCTTCAAGCCATGTTTATATCAGGATTTTCCGTGGTTTGTATGAAGAAAGAAACAG	2216
QY	2452	CACGTGAGACTTTGACAAACCAAGGCTTGCCAAATATGTCGGCTCAAGTTAAAGGGCTATAGG	2511
Db	2217	CACAGAGAGCTTGACAAACAATATTAGCCATGATATTACACAAATTCAAAGAGCAACAG	2276
QY	2512	TTTCCAGGCTGCTGTCATTTACCCACAATATTAGCAATCTTGGACAGGACATTTATATTC	2571
Db	2277	TTTCCAGGATTTGGCTGCTTTACACAAAAAGCAATACATGGGACTTTTCAGTTATCAATTC	2336
QY	2572	CTTAAATCTAATGTTGGCAATTTAACCTTTTACTCTGTAGCAATTTACCACATTCGTAAT	2631
Db	2337	CTTAAATCTAATGTTGGAGATGACATTTCTTATTTCTAGATTTTCTTCCAGTACTGCGGT	2396
QY	2632	ACCAGAGATGTTGTAATAGAAATATGTTGTCTGGACAAGCACTGAAGATTAAGAAGAGCT	2691
Db	2397	GACAGGATTTGATGAACCCGACGAATGACGTGATTTTGGCCAAACAAAGATTAACGAACAT	2456
QY	2692	AGAAAGAGCTGGGAATGCTGACAGAGCAATCGGAATTTCCGAACGTGTGTTCTTT	2751
Db	2457	TAAAGCATCTGGAAGATAGCAATGAAAGCTTTGGAGAAATATACGATATATGTGTCAAT	2516
QY	2752	GACTGGGAGCAAGATTTGTAATTCATGTAATGCAACAGATTTTGCACATGTCATTAAGAAA	2811

DB 2517 AACAGAGGAAAAAGCCTTCAGCAATGTATAGAGATGCTTCAGACTCAACACAGAAA 2576
OY 2812 CTCTTGAGGAAGAACACACATCTCGGGTCTCATTTTCTATACACCGGCAATGATGTA 2871
DB 2577 TACCTCGAAGAACACAGATATTTGGAAGCTGTATATGATTCACGATCGCTTATATA 2636
OY 2872 TTTTTCATATCGTGGCTGTTCCGGTTGTGTGCTACTGTTGGTGAAGTATGATTTGAA 2931
DB 2637 TTTTGCCATGACAGAGGTTTGATTTGAGAGCTTATTAATTCAGAGTGAACATGAC 2696
OY 2932 CTTTCAGAGTCTTTTGTATTCAGACTATTTGCTTTGGTGCATGAGCAGTGGGCA 2991
DB 2697 CCCAGAGGGCATGTTCAATAGTTTACTGCAATTCGCAATGAGCATGCGCATCGGAAA 2756
OY 2992 GGTGAGTTCAATTTGCTCTCTGACTATGCAAAAGCCAAAGTATCAGAGCCCGCATCAT 3051
DB 2757 AACGCTGTTTGGCTCTGTAATTTCCAAAGCCAAATCGGGGGCTGCGCATCTGTTTGC 2816
OY 3052 GATCATTTAAAAAAGCCCTCGATTTGACAGCTACAGCCCTCAGGCGCTCAAGCCAAATAC 3111
DB 2817 CTGTGTGAAAGAAACCAATATATAGACAGCCGACAGTCAAGAAAGGAAAAAGCAGACAC 2876
OY 3112 GTTGAAGAAATGTGACATTTAATGAGTCTGTCTCACTATCCACTGACAGACAGACAT 3171
DB 2877 ATGTGAAGGAATTTAGAGTTTCGAAAGTCTCTTCTTCTATCCATGATGCGCCAGATGT 2936
OY 3172 CCCCCTGCTCCAGAGGGCTGAGGCTCGAGGTGAGAAAGGCGCAGAGCGTGGCCCTGTAAG 3231
DB 2937 TTTTCATCTCTCGTGGCTTATTCCTCAGTATTTAGACGAGAAAGAAACATGACATTTGGG 2996
OY 3232 TAGCAGTGGCTGTGGGAAGACAGAGTGTTCAGTCTCTAGAGCGCTTCTATGACCCCT 3291
DB 2997 GAGCAGCGCTGTGGAAAAAGCACTTCTGTCAACTTCTGACAGACTTATATGACCCCT 3056
OY 3292 GCGTGGTACGTCTATTTGATGAGCAAAAGATTAAGCAACCTGCAATGCTCAGTGGCTCG 3351
DB 3057 GCAGGACAGAGTGTGTGATGAGTGTGATGATCAAAAGATTTGAATGTACAGTGGCTCG 3116
OY 3352 AGCAGACTGGGCACTGCTCTCAGAGAGCCATCTGTTTGTAGTGCAGCATTTGCCGAGAA 3411
DB 3117 TTTCCCAATATGCAATGCTTCTCAAGAGCTGTGCTTCAACTTCAACATTTGCTGAGAA 3176
OY 3412 CATTCCTATGAGACAAACAGCGGCTGTATACATGAGATTTATGACAGGACGACCA 3471
DB 3177 CATGCTATGATGACAAAGCGGTGTGCTGATTTAGATGATCAAGAAAGACGCGCAA 3236
OY 3472 GAGAGGCAACATACACACTTCATCGAGACACTCCCTGAGAAATACACACCAAGTATAG 3531
DB 3237 TGCAGCAAAATATCATTTCTTTATTTGAAGTCTCCCTGAGAAATACACACCAAGTATAG 3296
OY 3532 AGACAAAGAACCCAGCTCTCTGTGGCCAGAAAGCCATTTGCAATGCTCGGCTCT 3591
DB 3297 ACTAAAGAGACACAGCTTCTGTGGCCAGAAACAAAGACTGATGATTTGCAAGGCTCT 3356
OY 3592 TGTAGACAGCTCATATTTTGTGTTTGTATGATGATGATGATGATGATGATGATGATGAT 3651
DB 3357 TCTCAAAAACCCAAATTTTATTTATTTGATGATGATGATGATGATGATGATGATGATGAT 3416
OY 3652 TGAAGAGTTTGTCCAAAGAGCCCTGAGCAAAAGCCAGAAAGCCGACCTGATTTGTAT 3711
DB 3417 TGAGAAAGTGTGTCAGATGCTCTTGTATTAAGCCAGAGACGGGAAGACATGCTATGTGT 3476
OY 3712 CGCCGACGCTTGTCCACATTCAGATTCAGATTTAATATGATGATTTTCAGATGGCAA 3771
DB 3477 CACTCAGAGCTCTCTGTAATTCAGAAAGCAGATTTGATGATGATGATGATGATGATGATGAT 3536
OY 3772 AGTCAAGAGCAGTGCACACATCAACAGCTGCTGCGCCAGAAAGGCAATGATTTTTCAT 3831
DB 3537 GATTAAGGACAAAGAACTCATCAAGAGCTCTCTGAGAAATGAGACATATATTTTAAGTT 3596
OY 3832 GGTCAAGTGTCCAGGCTG 3848

DB 3597 AGTGAATGCACAGCTCAG 3613

RESULT 15
US-09-873-409-13
; Sequence 13: Application US/09873409
; Patent No. US20020037522A1
; GENERAL INFORMATION:
; APPLICANT: Frank, Markus
; APPLICANT: Sayegh, Mohamed
; TITLE OF INVENTION: A Gene Encoding a Multidrug Resistance Human P-Glycoprotein
; FILE REFERENCE: 81994/268611
; CURRENT APPLICATION NUMBER: US/09/873,409
; NUMBER OF SEQ ID NOS: 19
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 13
; LENGTH: 3702
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: NO. US20020037522A1e
; LOCATION: (723)..(723)
; OTHER INFORMATION: n at position 723 represents any nucleotide (A, T, C or G)
US-09-873-409-13
Query Match 28.8%; Score 1232.8; DB 10; Length 3702;
Best Local Similarity 61.2%; Pred. No. 1,9e-372;
Matches 2189; Conservative 0; Mismatches 1257; Indels 132; Gaps 7;
OY 358 CACGTATGCCATTTTACATGAGATCGGCTGCGGCTGCGGCTGCTGATCAATCA 417
DB 162 CAGTTTGAACCCGTATTTTGTGATATGATGATGATGATGATGATGATGATGATGATGAT 221
OY 418 GGTTCATTTCTGCTGCTGCGAGCAGGAAGACAGATCTCAATTTGAAAGAAATTTT 477
DB 222 GATTTCCTTGTGATTTAATTAATCAATGACAGACAGACAGACAGACAGACAGACAGACAG 281
OY 478 TCATGCTATCATGACAGAGAGATTTGCTGTTGAGCTGATGATGATGATGATGATGATGATGAT 537
DB 282 TCATGCTATGTTGGACAGAGACATGCTGTTGATGATGATGATGATGATGATGATGATGATGAT 341
OY 538 CACCGGCTCAGACAGAGATGCTCCAAATCAATGAAATGAAATGAAATGAAATGAAATGAAAT 597
DB 342 CACTGCATGACAGACATTTG---ACAATATGATGATGATGATGATGATGATGATGATGATGAT 398
OY 598 GTTCTTCACTCAATAGCAACATTTTTCACCGGTTTATGATGATGATGATGATGATGATGATGAT 657
DB 399 GTTGTTCAAAACATGCTACTTCTTTCATGATGATGATGATGATGATGATGATGATGATGATGAT 458
OY 658 GAAGCTAACCTTGTGATTTTGGCCATCAGCCCTGTTCTTGTGATGATGATGATGATGATGATGATGAT 717
DB 459 GAACTCAGCTAGTACATCTATCCACGCTCTCTTATATATGATGATGATGATGATGATGATGATGATGAT 518
OY 718 GGCAGAACTATCTTCAATTTACTATTAAGAACTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 777
DB 519 TTTAGAT 578
OY 778 AGTAGCTGAGAGAGTCTTACAGCAATCAGAACTGATTTGCTTTTGGAGAGCAAAAGAA 837
DB 579 TGTGAGAGAGAGTCTTGTTCATCAATCCGAAAGTCAATGATGATGATGATGATGATGATGATGATGAT 638
OY 838 AGAAC----- 842
DB 639 AGAATCTCAAGTCTTCTCTTTAATATATACAGATATGCTGTTTATTTTCCCA 698
OY 843 -----TTGAAGGTACAAACAAATTTAGAGAAAGCTAA 876
DB 699 GTGCTACTAAGTTGTGTTCTGTTTNTTGTATAGTATACAGAAATCTCAAAAGCTCAA 758
OY 877 AGGAATGGATTAAGAAAGCTATACGCGCAACATTTCTATTTGGTGGCGCTTTCTATT 936

Db 759 GGATTTGGCATAAAAAGACATAGCTTCAAAAGTGTCTGTGTGTGTACTCTT 818
QY 937 GATCTATGCATCATGCTGTGCTTGTGATGGAGCTCTCTGTCTCTCCAGTGA 996
Db 819 TATGATGGAACTATGACTTGTCTTTTGTATGAACTCTCTGATTTCTTATGAGA 878
QY 997 -----ATATCTATGAGACAAGTACTCATCTCTCTTTTGTATTAATGGGCTTT 1050
Db 879 AACTGGATATACATCGGAGCTGTCTTGCTTTCTTATGATATCATATGCACTTA 938
QY 1051 TAGATTTGACAGCATCCCCAACATTTGAAGCTTTGCAACGACAGAGACACTTA 1110
Db 939 TTGATTTGGACAGACAGTCCCTCACTTTGAAACCTTCGCATATGCCCCGAGAGCTCTT 998
QY 1111 TGAATCTTCAAGATTAATGACATTAACCAAGCATTTGACAGTATTTCCAGAGTGA 1170
Db 999 TCATATTTTCCAGGTATGATTAAGAAACCATATAGTAACCTTTTCCACAGCTGGATA 1058
QY 1171 TAAACGATATATTTAAGGAAATTTGGAATTCAAAAATTTCACTTCACTTACCTTC 1230
Db 1059 TAAACCTGAATCCATAGAGGAACGTGGAAATTTAAAAATGTTTCTTCAATATCCATC 1118
QY 1231 TCGAAAAAGAGTTAAGATCTTAAAGGCTCTCAACCTGAAGGTTCCAGAGTGGACAGT 1290
Db 1119 AAGACATCTATCAAGATTTCTGAAGGTCTGAATCTCAAAATTAAGTCTGGAGAGACAGT 1178
QY 1291 GGGGCTGTGGAGACAGTGGCTGGCGGAGACAGACGCTGCAGTGTATGACAGAGCT 1350
Db 1179 CGGCTGTGGTGGTCAATGAGAGAGTGGAGAGTACGGAGTCCAGTCTTCGAGAGGTT 1238
QY 1351 CTATGACCCACAGATGGATGGTCTGATTTGATGACAGACATTTAGACCATTAATGT 1410
Db 1239 ATATGATCCGAGATGGCTTATCATGTGTGATGATGAGATGCAATCAAGCTTTAAATGT 1298
QY 1411 AAGCATCTTGGGAAATTTACTGTGTGTGATGATGAGAGAGCTGTGTTGGCACAC 1470
Db 1299 GGGGCTTTTCAGAGCAATTAATGAGTGTATGATCAAGAGCTGTTTGTTCGGAGACAC 1358
QY 1471 GATAGCTGAACCATTCGTATGGCGGAGAAATGTCACATGATGATGATGAGAAAG 1530
Db 1359 CATCAGTAAATCATATCAATGATGAGAGATGATGTGATGATGATGATGAGAGAG 1418
QY 1531 TGTTAAGAGACCATGCTTATGATTTATCATGAACCTACCTAATTAATTTGACACTC 1590
Db 1419 ACGAAGGAGAGCAATGGTATGATTTATCATGAGATTCCTATTAATTTATATCAT 1478
QY 1591 GGTGGAGAGAGAGGGCGGAGTGAAGTGGAGACAGAAACAGATGCGCATTCG 1650
Db 1479 GGTAGGGAAAAAGAGAGTCAATGAGTGGAGGGCAGAGAAACAGAGATGCAATTCG 1538
QY 1651 GGCCGTGGTGGCAACCCCAAGATTTCTGTGATGAGAGCAACGTCAGCTCTGACAC 1710
Db 1539 TGCCTTATGATGAAACCCCAAGATTTGATTTAGATGAGGCTACAGTCTGCCCTGATTC 1598
QY 1711 TGAAGTGAAGCAGTGGTCAAGTGGCCGTGAATTAAGCCAGAAAAAGCCGAGTACAT 1770
Db 1599 AGAAAGCAAGTCAGCTGTTAAGCTGCATGGAGAGGGAGAGAAAGGTGGAGTACAT 1658
QY 1771 TGTATAGCTCATGTTGTCTACAGTGTGATGCGGATGTCATGCTGCTTTGATGA 1830
Db 1659 CGTGTGTGACACCGACTTCTCTATTTGCAAGTGTGATGATGATGATGATGATGAT 1718
QY 1831 TGGAGTCTGTGGAGAAAGGAATCATGATGAACCTGAAGAGAGAGGAGCTTACTT 1890
Db 1719 TGGAAAGCTGGCGAGAAAGAGCAATGCTGAACCTAATGGCAAAACGAGTCTATTA 1778
QY 1891 CAACTTGTACAAATGCAGACAGAGAAATGAATGATGATTAAGAAATGCCACTGTGA 1950
Db 1779 TTTACTTTGATGTGACAGATATTAATAAAGCTATGA-ACAGATGAGTCAATGACAT 1837
QY 1951 ATCCAAAAATGAAAGTATGCTTGGAAATGTCTCCAAAAATTCAGAGGTCCATTTAAT 2010
Db 1838 ATTCTACTAAAAAGAGCACTCATCTCTGCACTCTGTGAAG-----CAT 1889

QY 2011 AAAAAGATCAACTGACAGAGTATTCATGACCAACAAAGGCCAGACAGAAAGCTTGG 2070
Db 1890 CAAGTCAAGCTTCATGTCAGAAAGGCTGAGGAATCCACCAATCTAAGAGATAGTCT-- 1947
QY 2071 TACAAAAGAGACTGAATGAAATGACCCGAGCTTCTCTGAGAGGATTTGAGCT 2130
Db 1948 -----CCTGAAGTCTCTATTAATTAATTAATTAATTAATTAATTAATTAAT 1979
QY 2131 GAATCAACTGAATGAGCTTATTTTGTGGTGTATATTTGTCTATTTAATAAGGAG 2190
Db 1960 AAAAAAGGCTGAAGAGGCTTTTGTGGTGTGAGAGATTTGCTGTCTAATATGAAAC 2039
QY 2191 CCTCAACAGCATTTTTCATTAATTTTCAAGGATTTATGAGATTTTACCAGATGA 2250
Db 2040 TGTTCATTCAGATATTTTCCATCTCTTGGAAAAATTAATACCATGTTT---GGAATTA 2096
QY 2251 GGATCTGAAAACAAAACGACAGATAGTAACATGTTTCTGTATTTGTTCTGATCTGG 2310
Db 2097 TGATTAACCAACATTAAGCATGATCAGAAATTTATTCATGATATTCGATTTTGG 2156
QY 2311 AATTAATTTCTTATTAACATTTTCTCCAGGCTTCACATTTGGCAAAAGCTGGAGAT 2370
Db 2157 TGTATTTGCTTGTCTCATTTATTTCAATGCAAGGATTTATTTTCGGCAGAGAGGGAAT 2216
QY 2371 CCTCACTAAGCGGCTTCGATACATGTTTTCAGATTCATGCTGAGACAGAGATGTCAGT 2430
Db 2217 TTTTAAAGTATGATTAAGACATCTGGCCCTTAAACCATGTTATATTCATGATATTTGCC 2276
QY 2431 GTTTGATGACCTTAATAAACACACAGTGGAGATTTGCAACACAGGCTTGGCAATGATGGC 2490
Db 2277 GTTTGATGAAAAAGAAACAGACAGAGGCTTGCACAAATATATAGCATATATATAGC 2336
QY 2491 TCAAGTTAAAGGGCTATAGGTTCCAGGCTTGTCTGATTAACCCAGATATAGAAATCT 2550
Db 2337 ACAAATCAAGGAGCAAGAGTTCAGAGATGGGCTCTTAACAAAAATGCACTAATCAT 2396
QY 2551 TGGACAGCATATTAATATATCTTAATCTATGATGTTGGCAATTAACCTTTACTTAC 2610
Db 2397 GGGACTTTCAGTATTCATTTCTTATATATGATGATGAGATGATGATGATGATGATGAT 2456
QY 2611 AATTTACCATCATTCATGATAGAGAGTGTGAATGAATGAATGATGATGATGATGATGAT 2670
Db 2457 TATTCCTCAGTACTGCTGATGACAGGATGATGATGATGATGATGATGATGATGATGAT 2516
QY 2671 ACTGAAGATGAAGAAAGCTAGAGAGCTGGAGAGATTTGCTAAGAAAGCCATCGAAAA 2730
Db 2517 CAACAAGATGAAGCAAGAACTTAAGCATGCTGGAAGATGCAAGTGAAGCTTTGAGAA 2576
QY 2731 CTTCGAACTGTTTCTTGTGACTGCGAGAGAGAGTGTGAATGATGATGATGATGATGAT 2790
Db 2577 TATAGTACTATATGCTATTAACAGAGAAAAAGCTTTCGACAAATATATGAGAGAT 2636
QY 2791 TTTCAAGTACCATCAGAAAGCTTTGAGAGAGACACATCTTGGGGGTGCATTTTC 2850
Db 2637 GCTTCAACATCAACACAGAAATACCTCGAAGAGACAGATATTTGAGAGCTGTTATGC 2696
QY 2851 TATCAACAGGACATGATATTTTCTTATGCTGTGCTTTTCCGTTGGTGTGCTACTT 2910
Db 2697 ATTCAGCATGCTTTATATATTTTCTTATGCTGAGAGGTTTCATTTGAGAGCTATTT 2756
QY 2911 GTTGGCAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2970
Db 2757 AATTCAGGCTGAGCAATGACCCAGAGGCAATGTTTCTATGATTTTACTGCAATTTGATA 2816
QY 2971 TGGTGCATGGAGAGTGGGAGAGTCAATTTCTCTCTGACATATGCAAAAGCCAAAGT 3030
Db 2817 TGGAGCTATGGCATGCGAAAAAGCTGTTTGTGCTCTGAATATTCCTAAGCCAAATC 2876
QY 3031 ATTCAGGCCACGTCATCATGATCATTTGAATAAAAGCCCTGATTTGACAGCTACAGCC 3090
Db 2877 GGGGCTGGCATGCTTGTGCTTGTGAAAAAGAAACCAATATATGACAGAGCCGACGTCA 2936

OY	3091	TCACGGCCCTCAGCCAAATAGCTGTTGAAGCAAAATGTGACATTTAATGAGTGTGGTCAA	3150
Db	2937	AGNAGAAAAAACCAGACACATGCGAAGAATTTTAGACTTTTGAGATTCCTCTTTCTT	2996
OY	3151	CTATCCACTCCAGCACAGACATCCCCGTGCTCAGAGGGCTGAGGCTCTGAGGTMAAGAGG	3210
Db	2897	CTATCCATGTGCGCCAGATGTTTTATCTCTCCGTGGCTTATCCCTCAAGTATTGACGAGG	3056
OY	3211	CCAGAGCTGGCCCTGTAGTAGCATGTGGCTGTGGAGAGCAGCAGTTGTTACGCTCCT	3270
Db	3057	AAAACAGTAGCATTTGTGTGGGAGGAGCGGCTGTGGGAAAAAGCACTTGTCTCACTCT	3116
OY	3271	AGAGCGCTTCTATGACCCCTTGCTGTGTTCAGTGTCTAATTGATGGCAAGAGATTAAGA	3330
Db	3117	GCAGAGACTTTATGACCCCGTGTGAGGCAAAATGCTGTTGATGTGTGTGATGCCAAAAGA	3176
OY	3331	CCTGAATGTCCAGTGGCTCCGAGACACCTGGGCACTGCTGTCTGAGGAGCCATCCTGTT	3390
Db	3177	ATTGAATGTATAGAGGTCCGTTCCCAATTAGCATCTGTTCTTCAAGAGCTGTGCTCTT	3236
OY	3391	TGACCTGAGCATTTGSCGAGAACATTTGCTTATGAGAGACAACGCCGGTCTGTATCATATA	3450
Db	3237	CACCTGAGCATTTCTGTAGAAACATGCTATGTGTGACACAACGCGTGTGGTCAATTAGA	3296
OY	3451	AGAGATTATGACAGCAGCCAAGAGGCCCMAATPACACACTGTATCGAGACACTCCCTGA	3510
Db	3297	TGAGATTCAAGAACCCCAATGAGCAAAATATCCATTCTTTATTAGAGGTCTCCCTGA	3356
OY	3511	GAAATPACAACACAGAGTAGAGAGACAAGAGAACCCAGCTCTGTGGTGCAGAAACAGCG	3570
Db	3357	GAAATPACAACACAGTGTGACATGAAAGAGACACAGCTTTCTGTGGGCCCGAAACAAAG	3416
OY	3571	CATTGCCATAGCTGCGCTCTTGTTAGACACCTCATATTGTTGCTTTGGATGTAAGCTAC	3630
Db	3417	ACTAGCTATTGGCAAGGCTCTCTTCCAAAACCCAAATTTTATTGTTGATGAGGGCAC	3476
OY	3631	ATCAGCTGTGATACAGAAATGAGAAAGTTGTCCAAAGAGCCCTGACCAAGCCAGAGA	3690
Db	3477	TTTAGCCCTCGAATAATGACAGTAGGAAGGTGGTTCAAGCATGCCCTGTGATTAAGCCAGAC	3536
OY	3691	AGGCCACCTCATGTTGATGCGCCACCGCTGTCTCACATCOCGAATGACGATTATAT	3750
Db	3537	GGGAAGACATGCGTAGTGTGCTACACAGGCTCTCTGCATTTCAAGACGCAAGATTATAT	3596
OY	3751	AGTGTGTTTCAGAAATGCGAAAGTCAGAGACATGGCACATCAACAGCTCTGGCCCA	3810
Db	3597	AGTGTGTTCTCACAAATGAGAAAGATTAAGGAACCAAGAACTCATCAAGAGCTCTGAGAAA	3656
OY	3811	GAAAGCATATATTTTCCATGGTCAGTCCAGGCTG	3848
Db	3657	TCGAGACATATATTTTAAGTTAGTGAATGACACAGCTCAG	3694

Search completed: December 10, 2002, 05:54:15
Job time : 967 secs

GenCore version 5.1.3
Copyright (c) 1993 - 2002 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: December 9, 2002, 17:09:38 ; Search time 29467 Seconds

(without alignments)
2351.800 Million cell updates/sec

Title: US-09-672-725C-1

Perfect score: 4279

Sequence: 1 ggaagcagagatcgagatg.....caaaaaaaaaaaaaaa 4279

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 16154066 seqs, 8097743376 residues

Total number of hits satisfying chosen parameters: 32308132

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Database :

Listing first 45 summaries

EST:*
1: em_estba:*
2: em_esthum:*
3: em_estin:*
4: em_estnu:*
5: em_estcov:*
6: em_estcpl:*
7: em_estro:*
8: em_hlc:*
9: gb_est1:*
10: gb_est2:*
11: gb_hlc:*
12: gb_est3:*
13: gb_est4:*
14: gb_est5:*
15: em_estfun:*
16: em_estom:*
17: gb_gss:*
18: em_gss_hum:*
19: em_gss_inv:*
20: em_gss_pla:*
21: em_gss_vrt:*
22: em_gss_fun:*
23: em_gss_mam:*
24: em_gss_mus:*
25: em_gss_other:*
26: em_gss_pro:*
27: em_gss_tod:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	769.8	18.0	943	9	AL520322
2	738	17.2	929	14	BO882401
3	723.8	16.9	998	14	BM904842
4	651.4	15.2	861	14	BO720763
5	612.6	14.3	861	14	BO717101
6	576.8	13.5	981	12	BF969667

C	7	520.8	12.2	643	14	BO366012
	8	518.6	12.1	785	10	AV709991
	9	516.4	12.1	851	12	BF969062
	10	510.6	11.9	1019	12	BG248052
	11	509.6	11.9	894	12	BF584668
	12	479	11.2	803	12	BG298756
	13	479	11.2	587	10	BM053000
	14	475.2	11.1	605	10	AV616675
	15	467	10.9	795	12	BF333560
	16	467	10.9	563	13	BI964198
	17	467	10.9	563	13	BI964218
	18	463.8	10.8	563	13	BM052878
	19	449.4	10.5	563	13	BM053038
	20	449.4	10.5	944	12	BF796582
	21	448	10.5	760	9	AL520321
	22	438.4	10.2	726	12	BG293345
	23	436.4	10.2	546	13	BI964137
	24	417.4	9.8	662	13	BG963638
	25	408	9.5	759	10	BE376459
	26	407.4	9.5	517	13	BI964115
	27	386.2	9.0	598	9	AI807162
	28	386	9.0	1201	12	BE304587
	29	383.8	9.0	511	9	AA239727
	30	373.8	8.7	560	12	BE749379
	31	372	8.7	482	13	BI344244
	32	368.2	8.6	584	13	BI195085
	33	358	8.4	609	10	AM188349
	34	356	8.3	519	9	AA434959
	35	353.2	8.3	431	14	BO311985
	36	352.8	8.2	528	9	AA456377
	37	349.4	8.2	713	9	AI765301
	38	337.8	7.9	414	12	BE775493
	39	334.2	7.8	411	13	BM432014
	40	333.8	7.8	478	10	BE217416
	41	329	7.7	628	12	BG080311
	42	328.6	7.7	880	12	BC105127
	43	327.6	7.7	899	9	AV287606
	44	325.4	7.6	822	13	BI332761
	45	325	7.6	453	13	BM658768

ALIGNMENTS

RESULT 1
LOCUS AL520322 943 bp mRNA linear EST 13-FEB-2001
DEFINITION AL520322 LTI_NFL004_NBC2 Homo sapiens cDNA clone CS0DB006YCI15 5
ACCESSION AL520322
VERSION AL520322.1 GI:12783815
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 943)
AUTHORS Li, W.B., Gruber, C., Jesse, J. and Polayes, D.
TITLE Full-length cDNA libraries and normalization
JOURNALS Unpublished (2001)
COMMENT Contact: Genoscope
Genoscope - Centre National de Sequencage
BP 191 91006 EVRY cedex - France
Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr.

FEATURES

source
1..943
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="CS0DB006YCI15"
/clone_lib="LTI_NFL004_NBC2"
/sex="male"
/tissue_type="neuroblastoma cells"
/lab_host="DH10B"

/note="Organ: brain; Vector: PCMVSPORT 6; 1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime end enriched, double-stranded cDNA was digested with Not I and cloned into the Not I and Eco RV sites of the PCMVSPORT 6 vector. Library was normalized. Library was constructed by Life Technologies, a division of Invitrogen 9800 Medical Center Drive, Rockville, Maryland 20850, USA Fax: (1) 301 610 8371 Email: flangel@life.com URL: http://fulllength.invitrogen.com"

BASE COUNT 253 a 225 c 253 g 211 t 1 others
ORIGIN

Query Match 18.0%; Score 769.8; DB 9; Length 943;
Best Local Similarity 89.8%; Pred. No. 4.1e-123;
Matches 847; Conservative 1; Mismatches 93; Indels 2; Gaps 2;

QY 2944 TCTTTGGTATCTCAGCTATTGTCTTGGTCCATGCGAGGCGAGTCAGTTCATT 3003
DB 1 TCTGTATATTTTTCAGCTCTTGTCTTGTGTCATGCGCGCGGCGCAAGTCAGTTCATT 60
QY 3004 TGCCTCGAGATAGCCAAAGCAAGATATCAGACGCCAGTCATCATGATCATTTGAAA 3063
DB 61 TGTCTCTGACTATGCCAAAGCAAAATATCAGACGCCAGTCATCATGATCATTTGAAA 120
QY 3064 AAGCCCTCTGATGACAGCTACAGCCCTCAGCGCCCTCAAGCAATAGCTTGAAGAAA 3123
DB 121 AAGCCCTTGTATGACAGCTACAGCCAGGAGGCTTAATGCCGAACACTTGAAGAAA 180
QY 3124 TGTGCAATTAATGAGTCTGTTTCAACTATCCACTGACAGACATCCCGTCTCCA 3183
DB 181 TGTCACTATTTGATGATGATGATATCCACCCGACGACATCCAGTCTCTCCA 240
QY 3184 GGGGCTGAGCTCGAGGTGAAGAGGGCCAGAGCGCGCTGTAGTAGTAAAGTGGCTG 3243
DB 241 GGGAGTGAAGCTCGAGGTGAAGAGGGCCAGAGCGCTGTAGTAGTAAAGTGGCTG 300
QY 3244 TGGGAAGAGCAGTGTGTTCACTCTAGAGCGCTGTATGACCCCTTGGTGTCTCACT 3303
DB 301 TGGGAAGAGCAGTGTGTTCACTCTAGAGCGCTGTATGACCCCTTGGTGTCTCACT 360
QY 3304 GCTAATTTGATGCAAAAGATTAAGCACTGATGTCAGAGGCTCGAGACACCTGGG 3363
DB 361 GCTGCTGTGATGCAAAAGATTAAGCACTAATTTTCAAGTGGCTCGAGACACCTGGG 420
QY 3364 CATCGTGTCTCAGAGGCCATCTCTTGTACTGACATTTGCCAGAAATTTGCTATGG 3423
DB 421 CATCGTGTCTCAGAGGCCATCTCTTGTACTGACATTTGCTGAGAAATTTGCTATGG 480
QY 3424 AGACAAAGCGGGGTGTATCAATGAAGAGATTATGACAGCAGCAAGAGGCCAAT 3483
DB 481 AGACAAAGCGGGGTGTATCAATGAAGAGATTATGACAGCAGCAAGAGGCCAAT 540
QY 3484 ACACCACTTATGATGACAGCACTCCCTGAGAATATCAACACAGATGAGAGCAAGAGAC 3543
DB 541 ACATGCTTATGATGAGTCACTGCTTAATTAATATAGCACTAAAGAGAGCAAGAGAC 600
QY 3544 CCAGCTCTCTGTGGCCAGAAAAGAGGCAATGTCATAGCTGCGCTCTTGTAGACACC 3603
DB 601 TCACCTCTCTGTGGCCAGAAAAGAGGCAATGTCATAGCTGCGCTCTTGTAGACACC 660
QY 3604 TCATATTTTGGTGTGATGAGACTACATGAGCTGTGATGATGAGAAAGTGGTGT 3663
DB 661 TCATATTTTGGTGTGATGAGACTACATGAGCTGTGATGATGAGAAAGTGGTGT 720
QY 3664 CCAAGAGAGCCCTGAGCAAAAGCAGAGAGGCCACCTCATTTGATGAGCCACCGCTT 3723
DB 721 CCAAGAGAGCCCTGAGCAAAAGCAGAGAGGCCACCTCATTTGATGAGCCACCGCTT 780
QY 3724 GTCCACCATCCAGATGAGATTAATAGTG-GTGTTCAGATGAGCAAAAGTCAAGAGC 3782
DB 781 GTCCACCATCCAGATGAGATTAATAGTG-TGTTCAGATGAGCAAAAGTCAAGAGC 840

QY 3783 ATGGCACATCATCAGAGCTGCTGCGCCAGAAAGCATCTATTTTCCATGTCAGTGTCC 3842
DB 841 ATGGCACATCATCAGAGCTGCTGCGCCAGAAAGCATCTATTTTCAATGTCAGTGTCC 900
QY 3843 AGGCTGAGCAAAAGCGCTGTGATGAGTGTGATGAGTGT 3885
DB 901 AGGCTGAGCAAAAGCGC-AGTGAATCTGTGATGAGTGT 942

RESULT 2

BO882401 929 bp mRNA linear EST 16-AUG-2002

LOCUS AGENCOURT_8627902 NIH_MGC_43 Homo sapiens cDNA clone IMAGE:6291782
DEFINITION 5', mRNA sequence.

ACCESSION BO882401 GI:22274409
VERSION BO882401.1
KEYWORDS EST.

SOURCE

ORGANISM Homo sapiens
Eukaryote; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

REFERENCE 1 (bases 1 to 929)
NIH-MGC http://mgi.nci.nih.gov/
NATIONAL INSTITUTES OF HEALTH, MAMMALIAN GENE COLLECTION (MGC)
TITLE Unpublished (1999)
JOURNAL Contact: Robert Strausberg, Ph.D.
COMMENT Email: cgabs@emill.nih.gov
Tissue Procurement: ATCC

CDNA Library Preparation: Rubin Laboratory
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
DNA Sequencing by: Agencourt Bioscience Corporation
DNA Distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LNL at:
http://image.llnl.gov

Plate: LNCM2492 row: C column: 15
High quality sequence stop: 677.
Location/Qualifiers

1..929
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:6291782"
/clone_lib="NIH_MGC_43"
/tissue_type="normal pigmented retinal epithelium"
/lab_host="DH10B (phage-resistant) Site:1: xhoI: Site:2:
/note="Organ: eye; Vector: pORF7; Site:1: xhoI: Site:2:
EcoRI; CDNA made by oligo-dT priming. Directionally
cloned into EcoRI/XhoI sites using the following 5'
adaptor: GGCACGAC(G). Library constructed by Ling Hong
in the Laboratory of Gerald M. Rubin (University of
California, Berkeley) using ZAP-cDNA synthesis kit
(Stratagene) and Superscript II RT (Life Technologies).
Note: this is a NIH-MGC Library. 1 others

BASE COUNT 281 a 176 c 237 g 234 t 1 others
ORIGIN

Query Match 17.2%; Score 738; DB 14; Length 929;
Best Local Similarity 90.0%; Pred. No. 1.2e-117;
Matches 836; Conservative 0; Mismatches 86; Indels 7; Gaps 4;

QY 758 GCGTATGAAAAGGTGGAGAGTAGTGAAGAGTCTTAGCAATTCAGAACT-GTAT 816
DB 1 GCGTATGAAAAGGTGGAGAGTAGTGAAGAGTCTTAGCAATTCAGAACTGGTAT 60
QY 817 TGCCTTGGAGAGCAAAAGAAAGACTTGAAGGTACACAAAATTTGAGAGACTTA 876
DB 61 TGCCTTGGAGAGCAAAAGAAAGACTTGAAGGTACACAAAATTTGAGAGACTTA 120
QY 877 AGGATTTGGATTAAGAAAGTATACAGGCCCAATTTATAGGTGCTGCTTCTTAT 936
DB 121 AAGAAATTTGGATTAAGAAAGTATACAGGCCCAATTTATAGGTGCTGCTTCTTAT 180
QY 937 GATCATGATATATGCTGCTGCTTCTGATAGGACCTCTTGTGCTCTCCAGTGA 996
DB 937 GATCATGATATATGCTGCTGCTTCTGATAGGACCTCTTGTGCTCTCCAGTGA 996


```

Db 779 GAGTGTATTAACCTATCCACCCGACGACATCCAGGCTTCAGGGAGCTAGCCCTG 838
Oy 3197 GAGGTGAAGAAGGGCCAGACGCTGCCCTGCT AGTAGAGTGGCTGTGGGAAGGACG 3255
Db 839 GAGGTGAAGAAGGGCCAGACGCTGCCCTGCTGGGCGACAGTGGCTGGGAAGGACG 898
Oy 3256 AGT-TGTTCACTCTCTAGAGGCTTATGACCCCTTGCTGGTGTGAGTGAATGATG 3314
Db 899 AGTGGTCCACTCTCTGAGGCGGCTTACAAACCCCTTGCGAAGAAATGCTCTTGATG 958
Oy 3315 GCAAGAGATTAAGC 3329
Db 959 GCAAAAGATTAAGC 973

RESULT 4
BO720763 939 bp mRNA linear EST 16-JUL-2002
LOCUS BO720763
DEFINITION AGENCOURT_8233170 Lupski_dorsal_root_ganglion Homo sapiens cDNA
ACCESSION BO720763
VERSION BO720763.1 GI:21859660
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 939)
AUTHORS NIH-MGC http://mgi.nci.nih.gov/.
TITLES National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgabps-remail.nih.gov
Tissue Procurement: Dr. James R. Lupski
cDNA Library Preparation: Life Technologies, Inc.
DNA Sequencing by: Agencourt Bioscience Corporation
Cloned distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
plate: LLAM13573 row: 0 column: 22
high quality sequence stop: 646.
Location/Qualifiers
1. 939
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone_lib="Lupski_dorsal_root_ganglion"
/sex="male"
/tissue_type="dorsal root ganglia"
/dev_stage="adult, 36 yr"
/lab_host="DH10B"
/notes="Vector: PCMV-SPORT6 (Life Technologies); Site_1:
NotI; Site_2: SalI; cDNA made by oligo-dT priming.
Directionally cloned using the following adaptors:
5'-TCGACCCAGGCGTCG-3' and
5'-GACTAGTCTAAGATCAGGCGGCGCCCTG(15)-3'. Size selected >
1 kb for average insert length 1.7 kb. This is a primary
library, non-amplified. Library constructed by Life
Technologies and donated by J. Lupski, M.D./Ph.D. (Baylor
College of Medicine) and is available through Life
Technologies."
BASE COUNT 272 a 202 c 221 g 244 t
ORIGIN
Query Match 15.2%; Score 651.4; DB 14; Length 939;
Best Local Similarity 91.3%; Pred. No. 1.e-102;
Matches 713; Conservative 0; Mismatches 66; Indels 2; Gaps 2;
Oy 2335 GCAAAAGCTGGGAGATCTCTCACTAAGCGGCTGATCATGTTTTCAGATCATGCTGA 2414
Db 1 GCAAAAGCTGGGAGATCTCTCACTAAGCGGCTGATCATGTTTTCAGATCATGCTCA 60

```

```

Oy 2415 GACAGATGTCAGCTGTTTATGACCCCTAAACACCACTGAGCATGACACAGGC 2474
Db 61 GACAGATGTCAGCTGTTTATGACCTAAACACCACTGAGCATGACATACAGGC 120
Oy 2475 TTGCAATGATGCGGCTCAAGTTAAAGGGCTATAGGTTCCAGGCTTGCTGTCAATACC 2534
Db 121 TCGGCATGATGCTGCTCAAGTTAAAGGGCTATAGGTTCCAGGCTTGCTGTAAATTTCC 180
Oy 2535 AGAATATAGCAAACTTTGGGACAGGCACTTTATATCCCTTAATCTATGTTGGTCAATTA 2594
Db 181 AGAATATAGCAAACTTTGGGACAGGCACTTTATATCCCTTAATCTATGTTGGTCAATTA 240
Oy 2595 CACTTTTACCTCTTGAATTTACCACTTACCACTTACCACTTACCACTTACCACTTACCA 2654
Db 241 CACTTTTACCTCTTGAATTTACCACTTACCACTTACCACTTACCACTTACCACTTACCA 300
Oy 2655 TGTGTCGACACAGCACTGAAAGATAGAAAGAGCTAGAAAGAGCTGGGAATGCTA 2714
Db 301 TGTGTCGACACAGCACTGAAAGATAGAAAGAGCTAGAAAGAGCTGGGAATGCTA 360
Oy 2715 CAGAAAGCATGAAACCTTCCGAACCTGTTTCTTGTGACTCGGAGACAGAACTTGAAT 2774
Db 361 CAGAAAGCATGAAACCTTCCGAACCTGTTTCTTGTGACTCGGAGACAGAACTTGAAC 420
Oy 2775 ACATGTATGACACAGAGTTTGAATACCATACAGAACTTTTGAAGACACATCT 2834
Db 421 ACATGTATGACACAGAGTTTGAATACCATACAGAACTTTTGAAGACACATCT 480
Oy 2835 TCGGGGTCTCATTTTCTATACCCGAGCAATGATTTTCTATGCTGGCTGTTCC 2894
Db 481 TTGGATTTCAATTTTCTTCCACCGAGCAATGATTTTCTATGCTGGCTGTTCC 540
Oy 2895 GGTGTCCTACTTGTGTCGCAATGATGATGATGATGATGATGATGATGATGATGATGAT 2954
Db 541 GGTGTCCTACTTGTGTCGCAATGATGATGATGATGATGATGATGATGATGATGATGAT 600
Oy 2955 TCTCAGCTATGCTCTTGTGTCGCAATGATGATGATGATGATGATGATGATGATGATGAT 3014
Db 601 TTTACACTGTTGCTTGTGTCGCAATGATGATGATGATGATGATGATGATGATGATGATGAT 660
Oy 3015 ATG-CCAAAGCCAAAGTATCAGACGCCACGTCATGATGATGATGATGATGATGATGATGAT 3073
Db 661 ATG-CCAAAGCCAAAGTATCAGACGCCACGTCATGATGATGATGATGATGATGATGATGATGAT 720
Oy 3074 ATTGACAGCTACAGC-CTCAGCGCTCAAGCCAAATACGTTGGAAGAAATGAGACTT 3132
Db 721 ATTGACAGCTACAGC-CTCAGCGCTCAAGCCAAATACGTTGGAAGAAATGAGACTT 780
Oy 3133 T 3133
Db 781 T 781

RESULT 5
BO717101 861 bp mRNA linear EST 16-JUL-2002
LOCUS BO717101
DEFINITION AGENCOURT_8101135 Lupski_sympathetic_trunk Homo sapiens cDNA clone
IMAGE:6189786 5', mRNA sequence.
BO717101
ACCESSION BO717101
VERSION BO717101.1 GI:21855998
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 861)
AUTHORS NIH-MGC http://mgi.nci.nih.gov/.
TITLES National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgabps-remail.nih.gov
Tissue Procurement: Dr. James R. Lupski

```


cDNA Library Preparation: Life Technologies, Inc.
 cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LINTL)
 DNA Sequencing by: Agencourt Bioscience Corporation
 Clone distribution: MGC clone distribution information can be
 found through the I.M.A.G.E. Consortium/LINTL at:
<http://image.llnl.gov>
 plate: L1M13587 row: 1 column: 19
 High quality sequence stop: 603.

FEATURES

Location/Qualifiers
 1. 861

/organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clone="IMAGE:6189786"
 /clone_lib="Lupski_symphatic_c_trunk"
 /sex="male"
 /tissue_type="sympathetic trunk"
 /dev_stage="adult, 16 yr"
 /lab_host="DH10B"
 /note="Vector: PCW-SPORT6 (Life Technologies); Site_1:
 Note: Site_2: SalI; cDNA made by oligo-dT priming:
 directionally cloned using the following adaptors:
 5'-TCGACCCACGCGCCG-3' and
 5'-GACTAGTCTAGATCGACGGGCGCCCT(15)-3'. Size selected >
 1 kb for average insert length 1.9 kb. This is a primary
 library, non-amplified. Library constructed by Life
 Technologies and donated by J. Lupski, M.D./Ph.D. (Baylor
 College of Medicine); available through Life
 Technologies."
 BASE COUNT 236 a 191 c 216 g 218 t

Query Match 14.3%; Score 612.6; DB 14; Length 861;
 Best Local Similarity 88.3%; Pred. No. 5.4e-99;
 Matches 688; Conservative 0; Mismatches 89; Indels 2; Gaps 2;

QY 2586 GGCATTAACTTTTACTCTTGGCAATTTGACCATTCATTCAGTAATGACGAGGTTGTG 2645
 DB 1 GGCACCTAACACTGTTACTCTTAGCAATTTGACCATTCATTCAGTAATGACGAGGTTGTG 60
 QY 2646 AATGTAATGTTGTCGTGACAGACACACGAAAGTAAGAAAGAGCTAGGGA 2705
 DB 61 AATGTAATGTTGTCGTGACAGACACACGAAAGTAAGAAAGAGCTAGGGA 120
 QY 2706 AGATTGTACAGAGCCATCGAAACTTCGAACTGTTGTTCTTGTGACCTGGGACAGA 2765
 DB 121 AGATGCTACTGAGCAATAGAAACTTCGAACTGTTGTTCTTGTGACCTGGGACAGA 180
 QY 2766 AGTTGAATACATATATGACAGAGTTTGCAGTACATACAGAACTCTTGGAGGAAG 2825
 DB 181 AGTTGAATACATATATGACAGAGTTTGCAGTACATACAGAACTCTTGGAGGAAG 240
 QY 2826 CACACATCTCGGGGCTCTATTTTCTATCCAGGCAATGATGTTTTCCTATGCTG 2885
 DB 241 CACACATCTTGGAAATATACATTTTCTCTCAACCAGCAATGATGTTTTCCTATGCTG 300
 QY 2886 GCTGTTCCGTTTGGTCTCTACTTGTGCAAAATGATTCATGAACTTCAGSAGTTTC 2945
 DB 301 GATGTTCCGTTTGGTCTCTACTTGTGCAAAATGATTCATGAACTTCAGSAGTTTC 360
 QY 2946 TTTTGTATTTCTCAGTATTTGTTGTTGTCATGCGAGTGGGCGAGTTCATTTG 3005
 DB 361 TGTAGATTTTTCAGTCTGTTGTTGTTGTCATGCGAGTGGGCGAGTTCATTTG 420
 QY 3006 CTCTGACTATGCAAGCAAGATATGAGCCAGCTCATGATGATTCATGAAAAA 3065
 DB 421 CTCTGACTATGCAAGCAAGATATGAGCCAGCTCATGATGATTCATGAAAAA 480
 QY 3066 GCCCTGTGATGACAGTACAGCCTTCAGGCGCTTAAGCCAAATACGTTGAGGAATATG 3125
 DB 481 CCCCTGTGATGACAGTACAGCCTTCAGGCGCTTAAGCCAAATACGTTGAGGAATATG 540
 QY 3126 TGCATTTT AATGAGGTGTTTCAACTATCCACTGACAGCAATCCCGTCTCAG 3184
 DB 1 TGCATTTT AATGAGGTGTTTCAACTATCCACTGACAGCAATCCCGTCTCAG 3184

DB 541 TCACATTTGGTGAAGTTGATTCACATTCACACCCGACCGACATPCCAGTGTCTCAG 600
 QY 3185 GGGGTGAGCCCTCGAGGTGAGAGAGGCCACGCTGGCCCTCTAGTAGCAGTGGCTGT 3244
 DB 601 GGACTGAGCCCTGGAAGTGAAGAGGGGACAGCCCTGGCTCTGTGGGACAGAGTGGCTGT 660
 QY 3245 GGAAGAGCAGATGTTTCAAGCTCTCTAGAGCGCTCTCTATGACCCCTTGGCTGTCTAGTG 3304
 DB 661 GGAAGAGCAGATGTTTCAAGCTCTCTAGAGCGCTCTCTATGACCCCTTGGGAGGAATG 720
 QY 3305 CTAATTATGCGAAGAGATTAAGCACCTGAATGTCCAGTGTCCGACACACCTGGG 3363
 DB 721 CTGCTGATGCGAAA AATTAAGCAGCTGAAGGTTCAATTGGCTCGAAGCCCTCTGG 778

RESULT 6
 BF969667 981 bp mRNA linear EST 22-JAN-2001
 LOCUS 602272046F1 NIH_MGC_84 Homo sapiens cDNA clone IMAGE:4360090 5',
 DEFINITION mRNA sequence.
 ACCESSION BF969667
 VERSION BF969667.1 GI:12336882
 KEYWORDS EST.
 SOURCE human.
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 REFERENCE NIH-MGC <http://mgc.nci.nih.gov/>.
 1 (bases 1 to 981)
 AUTHORS National Institutes of Health, Mammalian Gene Collection (MGC)
 TITLE Unpublished (1999)
 JOURNAL Contact: Robert Strausberg, Ph.D.
 COMMENT Email: cgabbs-remail.nih.gov
 Tissue Procurement: ATCC
 cDNA Library Preparation: Life Technologies, Inc.
 DNA Library Arrayed by: The I.M.A.G.E. Consortium (LINTL)
 Clone distribution: MGC clone distribution information can be
 found through the I.M.A.G.E. Consortium/LINTL at:
<http://image.llnl.gov>
 plate: L1M10000 row: 1 column: 11
 High quality sequence stop: 670.

FEATURES

Location/Qualifiers
 1. 981

/organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clone="IMAGE:4360090"
 /clone_lib="NIH_MGC_84"
 /tissue_type="adrenal cortex carcinoma, cell line"
 /lab_host="DH10B (phage-resistant)"
 /note="Organ: adrenal gland; Vector: PCW-SPORT6; Site_1:
 Note: Site_2: SalI; Cloned unidirectionally; oligo-dT
 primed. Average insert size 1.229 kb. Library enriched for
 full-length clones and constructed by Life Technologies.
 Note: this is a NIH-MGC Library."

BASE COUNT 267 a 194 c 250 g 270 t

Query Match 13.5%; Score 576.8; DB 12; Length 981;
 Best Local Similarity 92.1%; Pred. No. 7.6e-90;
 Matches 630; Conservative 0; Mismatches 52; Indels 2; Gaps 2;

QY 2393 ATGGTTTTCAGATCCATGCTGAGCAGAGTGCACCTGTTGATGACCTTAAGAACAC 2452
 DB 1 ATGGTTTTCAGATCCATGCTGAGCAGAGTGCACCTGTTGATGACCTTAAGAACAC 60
 QY 2453 ACTGAGAGCTTACACACAGGCTTCCCAATGATGGGCTCAAGTTAAAGGGCTATAGT 2512
 DB 61 ACTGAGAGCTTACACACAGGCTTCCCAATGATGGGCTCAAGTTAAAGGGCTATAGT 120
 QY 2513 TCCAGGCTTGTGCTATTAACCAAGATATACCAATCTTGGACAGCATTTATATCC 2572
 DB 121 TCCAGGCTTGTGCTATTAACCAAGATATACCAATCTTGGACAGCATTTATATCC 180

QY 2573 TTAATCTATGCTGGCAATTACACTTTTACTCTTACAGATTTGATCCCATTCATGCAATA 2632
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 Db 181 TTCTATCTATGCTGGCAACTAACACTGTTTACTTACAGATTTGATCCCATTCATGCAATA 240
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 QY 2633 GCAGAGATTTGTAATGAAATGTTGTCTGCAACAGCACTGAAGAATAGAAAGAGCTA 2692
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 Db 241 GCAGAGATTTGTAATGAAATGTTGTCTGCAACAGCACTGAAGAATAGAAAGAGACTA 300
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 QY 2693 GAAGAGCTGGGAAGATGCTACTGAGCAATAGAAACTTCCGAACCTGTTCTTTG 2752
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 Db 301 GAAGAGCTGGGAAGATGCTACTGAGCAATAGAAACTTCCGAACCTGTTCTTTG 360
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 QY 2753 ACTCGGAGCAGAGATTTGTAATGATGATGCAACAGATTTGCAAGTACCATACAGAAAC 2812
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 Db 361 ACTCGGAGCAGAGATTTGTAATGATGATGCAACAGATTTGCAAGTACCATACAGAAAC 420
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 QY 2813 TCTTTGAGGAAGACACACTTCCGGGCTCTCATTTTCTATCACCAGCAATGATGAT 2872
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 Db 421 TCTTTGAGGAAGACACACTTCCGGGCTCTCATTTTCTATCACCAGCAATGATGAT 480
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 QY 2873 TTTTCCATGCTGCTGCTGTTTCCGGTGTGCTACTTGTGCAATGATGATGATGATGAT 2932
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 Db 481 TTTTCCATGCTGCTGCTGTTTCCGGTGTGCTACTTGTGCAATGATGATGATGATGAT 540
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 QY 2933 TTTTCCATGCTGCTGCTGTTTCCGGTGTGCTACTTGTGCAATGATGATGATGATGAT 2992
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 Db 541 TTTTCCATGCTGCTGCTGTTTCCGGTGTGCTACTTGTGCAATGATGATGATGATGAT 600
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 QY 2993 GTCACTTATTTGCTCTGCTACTGATGCCAAGCAAGATGATGAGCCCACTGATGATG 3052
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 Db 601 GTCACTTATTTGCTCTGCTACTGATGCCAAGCAAGATGATGAGCCCACTGATGATG 658
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 QY 3053 ATCATTTGAAAAAGCCCTCTGATT 3076
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 Db 659 ATCATTTGAAAAAGCCCTCTGATT 682
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 RESULT 7
 BQ366012/c 643 bp mRNA linear EST 21-MAY-2002
 LOCUS BQ366012 260900-025-h11 GN0030 Homo sapiens cDNA, mRNA sequence.
 DEFINITION BQ366012
 ACCESSION BQ366012.1 GI:21041524
 VERSION
 KEYWORDS
 SOURCE
 ORGANISM
 Homo sapiens
 human.
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
 1 (bases 1 to 643)
 Dias Neto, E., Garcia Correa, R., Verjovski-Almeida, S., Briones, M.R.,
 Nagai, M.A., da Silva, W. Jr., Zago, M.A., Bordini, S., Costa, F.F.,
 Goldman, G.H., Carvalho, A.F., Matsukuma, A., Bala, G.S., Simpson, D.H.,
 Brugnabein, A., deoliveira, P.S., Bucher, P., Jongeneel, C.V., O'Hare
 M.J., Soares, F., Brentani, R.R., Reis, L.F., de Souza, S.J. and
 Simpson, A.J.
 Shotgun sequencing of the human transcriptome with ORF expressed
 sequence tags
 Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)
 20202663
 Contact: Simpson A.J.G.
 Laboratory of Cancer Genetics
 Ludwig Institute for Cancer Research
 Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-Sp,
 Brazil
 Tel: +55-11-2704922
 Fax: +55-11-2707001
 Email: asimpson@ludwig.org.br
 This sequence was derived from the FAPESP/LICR Human Cancer Genome
 Project. This entry can be seen in the following URL
 (http://www.ludwig.org.br/scripts/gethtml2.pl?pl=MR2&t2=MR2-GN0030-
 260900-025-h11&t3=2000-09-26&t4=1)
 Seq primer: puc 18 forward

FEATURES
 source
 High quality sequence start: 20
 High quality sequence stop: 641.
 Location/Qualifiers
 1..643
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clone_lib="GN0030"
 /dev_stage="Adult"
 /note="Organ: Placenta, normal; Vector: puc18; Site:1: SmaI
 ; Site:2: SmaI; A mini-library was made by cloning
 products derived from ORESTES PCR (U.S. Letters Patent
 Application No. 196,716 - Ludwig Institute for Cancer
 Research) profiles into the pUC 18 vector. Reverse
 transcription of tissue mRNA and cDNA amplification were
 performed under low stringency conditions."
 BASE COUNT 200 a 133 c 140 g 170 t
 ORIGIN
 Query Match 12.2%; Score 520.8; DB 14; Length 643;
 Best Local Similarity 88.7%; Pred. No. 4.2e-80;
 Matches 564; Conservative 0; Mismatches 72; Indels 0; Gaps 0;
 QY 2041 TGACACAAAGGCCAAGACAGAAAGCTTGTACAAAGAGAGCTTGATGAGATGTACC 2100
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 Db 642 TGATCAACAAGCCCAAGACAGAAAGCTTGTACAAAGAGAGCTTGATGAGATGTACC 583
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 QY 2101 TCCAGTTTCTTCTGGAGGATTCGAGCTGAGCTGAACCTGAACCTGAATGCTTTGTTGT 2160
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 Db 582 TCCAGTTTCTTCTGGAGGATTCGAGCTGAGCTGAACCTGAACCTGAATGCTTTGTTGT 523
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 QY 2161 TGTATATTTTGTGCTATATTAACGAGAGCCGCAACGCAATTTTCAATATATTTTC 2220
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 Db 522 TGTGTATTTTGTGCTATATTAACGAGAGCCGCAACGCAATTTTCAATATATTTTC 463
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 QY 2221 AAGGATTAATGAGGATCTTACCCGAGATGAGGATCTGAACCAACAAAGCAGATGTAA 2280
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 Db 462 CAAGATTAATGAGGATTTTACCAAGATGATGATGATGATGATGATGATGATGATG 403
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 QY 2281 CAGGTTTCTGATATGTTGTTTCTGATGCTTGGATATATTTCTTTATTCATTTTCTCCA 2340
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 Db 402 CTTGTTTCTGATATGTTGTTTCTGATGCTTGGATATATTTCTTTATTCATTTTCTCCA 343
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 QY 2341 GGGCTTCAATTTGGCAAGCTGGGAGATGCTGATGAGCGGCTTGATGATGATGATG 2400
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 Db 342 GGGCTTCAATTTGGCAAGCTGGGAGATGCTGATGAGCGGCTTGATGATGATGATG 283
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 QY 2401 CAGATCCATGCTGAGACAGAGATGCTGCTGTTGATGATGATGATGATGATGATGATG 2460
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 Db 282 CCGATCCATGCTGAGACAGAGATGCTGCTGTTGATGATGATGATGATGATGATGATG 223
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 QY 2461 ATTGACAAACGAGCTTGGCAATGATGCTGCTGATGATGATGATGATGATGATGATG 2520
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 Db 222 ATTGACAAACGAGCTTGGCAATGATGCTGCTGATGATGATGATGATGATGATGATG 163
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 QY 2521 TGGTGTATTAACCAAGATTAAGCAATTTTGGACAGCATTAATATATTCCTTAATCTA 2580
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 Db 162 TGGTGTATTAACCAAGATTAAGCAATTTTGGACAGCATTAATATATTCCTTAATCTA 103
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 QY 2581 TGGTGTATTAACCAATTTTACTTACTTACTTACTTACTTACTTACTTACTTACTTACT 2640
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 Db 102 TGGTGTATTAACCAATTTTACTTACTTACTTACTTACTTACTTACTTACTTACTTACT 43
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 QY 2641 TGTGAATGAAATGTTGTCTGAGACAGCACTGAA 2676
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 Db 42 TGTTCATTAATAATGTTGTCTGAGACAGCACTGAA 7
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 RESULT 8
 AV709991 785 bp mRNA linear EST 09-OCT-2000
 LOCUS AV709991
 DEFINITION AV709991 Cu Homo sapiens cDNA clone CUAJA09 5', mRNA sequence.
 ACCESSION AV709991
 VERSION AV709991.1 GI:10728272

||||| 181 CACGAAAGTAAGAAAGACTGAGAGGTCGGAGAGATCGCTACTGTAAGCAATAGAAA 240
OY 2730 ACTTCGCAAGTGTGTTCTTTTCTTCTGAGAGAGAGATTGTAATCATGTATGACAGA 2789
Db 241 ACTTCGCAAGTGTGTTCTTTTCTTCTGAGAGAGAGATTGTAATCATGTATGACAGA 300
OY 2790 GTTTCGCAAGTGTGTTCTTTTCTTCTGAGAGAGAGATTGTAATCATGTATGACAGA 2849
Db 301 GTTTCGCAAGTGTGTTCTTTTCTTCTGAGAGAGAGATTGTAATCATGTATGACAGA 360
OY 2850 CTATCAGCAGCAGCATGATGATTTTCTTCTGAGAGAGAGATTGTAATCATGTATGACAGA 2909
Db 361 CCTTCAGCAGCAGCATGATGATTTTCTTCTGAGAGAGAGATTGTAATCATGTATGACAGA 420
OY 2910 TGTTCGCAAGTGTGTTCTTTTCTTCTGAGAGAGAGATTGTAATCATGTATGACAGA 2969
Db 421 TGTTCGCAAGTGTGTTCTTTTCTTCTGAGAGAGAGATTGTAATCATGTATGACAGA 480
OY 2970 TGTTCGCAAGTGTGTTCTTTTCTTCTGAGAGAGAGATTGTAATCATGTATGACAGA 3028
Db 481 TGTTCGCAAGTGTGTTCTTTTCTTCTGAGAGAGAGATTGTAATCATGTATGACAGA 540
OY 3029 GTATCAGCAGCAGCATGATGATTTTCTTCTGAGAGAGAGATTGTAATCATGTATGACAGA 3088
Db 541 ATATCAGCAGCAGCATGATGATTTTCTTCTGAGAGAGAGATTGTAATCATGTATGACAGA 599
OY 3089 CCTTCAGCAGCAGCATGATGATTTTCTTCTGAGAGAGAGATTGTAATCATGTATGACAGA 3145
Db 600 ACGAAGGCGCTAATGTCGCGGAGACATTTGGAGAGAAATGTCACATGTGTGAAAGCTT 659
OY 3146 TTCACTATCCAGTC-----GACCAAGACATCCCGTGTCTCCAGAGGCGTCAAGCTCGA 3198
Db 660 GTTTTCAATTAATTCACACCCGTAGCGAGATCCCAAGTGTCTCCAGAGCTGAAGCTGG 719
OY 3199 GGTGAAGAAGGCGCAGAGCGTGGCCCTCG 3227
Db 720 GTGAAGAAGGCGCAGAGCGGCTCG 748

RESULT 10
Bg248052 1019 bp mRNA linear EST 13-FEB-2001
LOCUS 60235987F1 NCL_CGAP_Mam1 Mus musculus cDNA clone IMAGE:4488404 5',
DEFINITION mRNA sequence.
ACCESSION Bg248052
VERSION Bg248052.1 GI:12757867
KEYWORDS EST.
SOURCE house mouse.
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE NIH-MGC http://mgc.nci.nih.gov/
AUTHORS National Institutes of Health, Mammalian Gene Collection (MGC)
TITLE Unpublished (1999)
JOURNAL Contact: Robert Strausberg, Ph.D.
COMMENT Email: egabbs-remail.nih.gov
Tissue Procurement: Gilbert Smith, Ph.D.
cDNA Library Preparation: Life Technologies, Inc.
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
plate: LLAM10334 row: n column: 21
High quality sequence stop: 650.
location/Qualifiers
1. 1019
/organism="Mus musculus"
/strain="FVB/N"
/db_xref="taxon:10090"
/clone="IMAGE:4488404"

BASE COUNT 276 a 262 c 283 g 196 t 2 others
ORIGIN
Query Match 12.1%; Score 516.4; DB 12; Length 1019;
Best Local Similarity 83.5%; Pred. No. 26-79;
Matches 611; Conservative 0; Mismatches 116; Indels 5; Gaps 2;
OY 3113 TTGGAAGAAATGTGACATTAATGAGGCGTTCACATATCCACCTGACAGATC 3172
Db 1 TTGGAAGAAATGTGACATTAATGAGGCGTTCACATATCCACCTGACAGATC 60
OY 3173 CCCGCTCCAGAGGCGTGAAGTGAAGAGGCGCAGACCTGCGCTCGTAGGT 3232
Db 61 CCAGTGTTCAGAGGCGTGAAGTGAAGAGGCGCAGACGCTGCGCTCGTAGGT 120
OY 3233 AGCAGTGTTCAGAGGCGTGAAGTGAAGAGGCGCAGACCTGCGCTCGTAGGT 3292
Db 121 AGCAGTGTTCAGAGGCGTGAAGTGAAGAGGCGCAGACCTGCGCTCGTAGGT 180
OY 3293 GCTGTGTAGTGTGATTAATGAGGCGTGAAGTGAAGAGGCGCAGACCTGCGCTCGTA 3352
Db 181 GCTGTGTAGTGTGATTAATGAGGCGTGAAGTGAAGAGGCGCAGACCTGCGCTCGTA 240
OY 3353 GCACAGCTGGGCGTGTGTCTCAGAGGCGCAGACCTGCGCTCGTAGGTGCGAGAAC 3412
Db 241 GCACAGCTGGGCGTGTGTCTCAGAGGCGCAGACCTGCGCTCGTAGGTGCGAGAAC 300
OY 3413 ATTGCTTAATGAGGCGTGAAGTGAAGAGGCGCAGACCTGCGCTCGTAGGTGCGAGAAC 3472
Db 301 ATTGCTTAATGAGGCGTGAAGTGAAGAGGCGCAGACCTGCGCTCGTAGGTGCGAGAAC 360
OY 3473 GAGGCGCAATACACCACTTCATGAGACACTCCCTGAGAAATACACCAAGTAGTGA 3532
Db 361 GAGGCGCAATACACCACTTCATGAGACACTCCCTGAGAAATACACCAAGTAGTGA 420
OY 3533 GACAAAGAACCCAGCTGCTGAGGCGCAGAAACGCGCATTCGATGCGCTCTT 3592
Db 421 GACAAAGAACCCAGCTGCTGAGGCGCAGAAACGCGCATTCGATGCGCTCTT 480
OY 3593 GTTAGACAGGCTGATTTTGTGTTGATGAGTACATGAGTATGACAGGCGCTCTT 3652
Db 481 GTTAGACAGGCTGATTTTGTGTTGATGAGTACATGAGTATGACAGGCGCTCTT 540
OY 3653 GAAAAGTTGTCGAAGAGCCCTGAGCAAGGCGCAGAAAGGCGCAGCTGATGATGATC 3712
Db 541 GAAAAGTTGTCGAAGAGGCGTGAAGGCGCAGAAAGGCGCAGCTGATGATGATC 600
OY 3713 GCCACCGCTGTGTCACATCATGAAATGCAATTTAATG--TGTGTTTCAGAAATGCA 3770
Db 601 GCTACCGCTGTGTCACATCATGAAATGCAATTTAATG--TGTGTTTCAGAAATGCA 660
OY 3771 AAGTCAAGAGCATGACACATCAACAGC--TGTGTTTCAGAAATGCA 3827
Db 661 AAGTCAAGAGCATGACACATCAACAGC--TGTGTTTCAGAAATGCA 720
OY 3828 CCATGCTCAGTG 3839
Db 721 CCATGCTCAGTG 732

RESULT 11
Bf584668 894 bp mRNA linear EST 12-DEC-2000
LOCUS 602098406F1 NCL_CGAP_Co24 Mus musculus cDNA clone IMAGE:4218385 5',
DEFINITION mRNA sequence.

ACCESSION BF584668
 VERSION BF584668.1 GI:11658386
 KEYWORDS EST.
 SOURCE house mouse.
 ORGANISM Mus musculus

REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 1 (bases 1 to 894)
 NIH-MGC <http://mgi.nci.nih.gov/>
 TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
 JOURNAL Unpublished (1999)
 COMMENT Contact: Robert Strausberg, Ph.D.
 Email: cgabbs-remail.nih.gov
 Tissue Procurement: Jeffrey E. Green, M.D.
 cDNA Library Preparation: Life Technologies, Inc.
 DNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
 Clone distribution: MGC clone distribution information can be
 found through the I.M.A.G.E. Consortium/LLNL at:
<http://image.llnl.gov>
 Plate: LHAM9798 row: 1 column: 02
 High quality sequence stop: 651.
 Location/Qualifiers

FEATURES
 source
 1..894
 /organism="Mus musculus"
 /strain="FVB/N"
 /db_xref="taxon:10090"
 /clone="IMAGE:4218385"
 /clone_lib="NCI CGAP Co24"
 /lab_host="DH10B (T1 phage-resistant)"
 /note="Organ: colon; Vector: PCMV-SPORT6; Site_1: NotI;
 Site_2: SalI; Cloned unidirectionally. Primer: Oligo dT.
 Average insert size 1.6 kb. Constructed by Life
 Technologies. Note: this is a NCI CGAP Library."
 BASE COUNT 263 a 203 c 242 g 185 t 1 others
 ORIGIN

Query Match 11.9%; Score 510.6; DB 12; Length 894;
 Best Local Similarity 83.0%; Pred. No. 2.1e-78;
 Matches 629; Conservative 0; Mismatches 125; Indels 4; Gaps 4;

OY 1018 CACTGCTCTTTTCTGATTAATTTGGGGCTTTAGTATGACAGAGCCCAAGCAT 1077
 DB 1 CACTGCTCTTTTCCGGTTAA-TGGAGCATTCAGTGTGGACAGCATCTCAATAT 59
 OY 1078 TGAAGCATTTGCAACGAGAGGAGGAGGAGCTTATGAATCTTCAAGTAATTAACAATA 1137
 DB 60 TGAAGCCTTCCCAATGACAGAGGAGGAGGAGCTTATGAAGTCTTCAAAATATTAATA 119
 OY 1138 ACCAAGCTTACAGCTATTTGAGAGTGGACATTAACAGATTAATTAAGGAATTT 1197
 DB 120 GCCCAGTATACAGCTTCTCAAGAGTGGGCAAAACGACACATACAGGAATCT 179
 OY 1198 GGAATTCAAAAATGTACCTTACCTTTTCGAAAAGAGTTAGATCTTAAAGG 1257
 DB 180 GGAATTTAGAAATTTACTTCACTTCACTTCCATCTCGAAAAAGAGTTACATCTTGAAGG 239
 OY 1258 TCTCAACCTGAAGGTTGAGAGTGGGAGACAGAGGCGCTGGTGGAGAAAGTGGCGG 1317
 DB 240 CCTCAATCTGAAGGTGAAGAGCGGAGAGGCGCTGGTGGAGAAAGTGGCGG 299
 OY 1318 GAAGAGACACACCGTGCAGCTGATGAGAGGCTTATGAAGCCACAGATGAGCATGTCTG 1377
 DB 300 AAAAAGCACAACCTGTCAGCTGATGCAAGGCTTACAGCCCTTATGAGTGTGAG 359
 OY 1378 TATTTGATGACAGACATTAGGACCTTAATTAAGGCACTTTCGGAAATTAAGTGTGT 1437
 DB 360 TATTCAGCGGACAGACATCAGAAACATCATGTGAGTATCTGAGAGGAGATCTTGGTGT 419
 OY 1438 GGTGATCTGAGAGCTGTGTGTTTGCACACAGATAGCTGAAACATTCGCTTGGGCG 1497
 DB 420 GGTGATCTGAGAGCTGTGTGTTTGCACACAGATAGCTGAAACATTCGCTTGGGCG 479

OY 1498 GAAATATGCACCATGATGATGAGAAAAGCTGTTAAGCAAGCCCAATGCTATGATTT 1557
 DB 480 AGAATATGTCACCATGATGATGAGAAAAGCTGTCACAGAGCCCAATGCTATGATTT 539
 OY 1558 TATCATGAACCTACCTATATAATTTGACACTGTGTTGAGAGAGAGGGGCCAGCTGAG 1617
 DB 540 CATCATGAACCTGCCCCACCAATTTTGACACCTGTTGGTGGAGAGAGGGCGAGCTGAG 599
 OY 1618 TGTGTGACAGAAACAGAAATTCGCAATTCGCGGCCCTGGTGTGCGAACCCCAAGATTC 1677
 DB 600 TGGGGACAGAAACAGAAATTCG-CATTTGCCCGGCGCTGGTCCCAATCCCAAG-TGCC 657
 OY 1678 TCTGCTGATGAGGCAACGTCAGCTCTGACACTGGAAGTGAAGCAAGTGTTCAGGTGCG 1737
 DB 658 TTTGTGTGAGAGAGCCACCTTCAGC-CTGATATACAGAAAGTGAAGCTGGGGTTCAGGCGCG 716
 OY 1738 CTTGATTAAGCCCAAGAAAGCCGAGCTACCATTTGTA 1775
 DB 717 ACCTAGTATAGGTATCAAAAGGGGCGAGCACCATTTGGGA 754

RESULT 12
 LOCUS BG298756
 DEFINITION 60239681F1 NIH_MGC_94 Mus musculus cDNA clone IMAGE:451157 5',
 mRNA sequence.
 ACCESSION BG298756
 VERSION BG298756.1 GI:13063728
 KEYWORDS EST.
 SOURCE house mouse.
 ORGANISM Mus musculus

REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 1 (bases 1 to 803)
 NIH-MGC <http://mgi.nci.nih.gov/>
 TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
 JOURNAL Unpublished (1999)
 COMMENT Contact: Robert Strausberg, Ph.D.
 Email: cgabbs-remail.nih.gov
 Tissue Procurement: The Cepko Laboratory
 cDNA Library Preparation: Life Technologies, Inc.
 DNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
 Clone distribution: MGC clone distribution information can be
 found through the I.M.A.G.E. Consortium/LLNL at:
<http://image.llnl.gov>
 Plate: LHAM10394 row: b column: 22
 High quality sequence stop: 698.
 Location/Qualifiers

FEATURES
 source
 1..803
 /organism="Mus musculus"
 /db_xref="taxon:10090"
 /clone="IMAGE:451157"
 /clone_lib="NIH_MGC_94"
 /tissue_type="retina"
 /lab_host="DH10B (phage-resistant)"
 /note="Organ: eye; Vector: PCMV-SPORT6; Site_1: NotI;
 Site_2: SalI; Cloned unidirectionally; Oligo-dT primed.
 Average insert size 3.3 kb. Library enriched for
 full-length clones and constructed by Life Technologies.
 Note: this is a NIH-MGC Library."
 BASE COUNT 240 a 162 c 205 g 196 t

Query Match 11.9%; Score 509.6; DB 12; Length 803;
 Best Local Similarity 80.8%; Pred. No. 3.2e-78;
 Matches 643; Conservative 0; Mismatches 149; Indels 4; Gaps 4;

OY 729 TATCTTCATTTACTGATTAAGAACTCTTGCCCTATGCAAAAGCTGGAGCACTAGCTAG 788
 DB 10 TGTCTTCATTTACTGATTAAGAACTCATGCTATGCAAAAGCTGGAGCACTAGCTAG 69
 OY 789 AAGTCTTACACATTCAGAACTGTGATTCCTTTGAGAGCAAAAGAAAGAACTTGA 848

QY 3282 ATGACCCCTGGCTGCTCAATGCTAATGATGCGAAGATGAAGACACCATGATGTC 3341
 DB 421 AGACCCCTTGGCAGGGAAGTGTCTGATGACCAAGAAATGAAGCCAGTGAATTC 480
 QY 3342 AGTGGCTCCGAGCACCACCTGGGCATGCTGTCTCAGAGAGCCCATCTGTTGATCGAGCA 3401
 DB 481 AGTGGCTCCGAGCACCACCTGGGCATGCTGTCTCAGAGAGCCCATCTGTTGATCGAGCA 3401
 QY 3402 TTGGCGAAGACATGCTGATGAGACAGACCGGG 3437
 DB 541 TTGCTGAGAACATTTGCTATGAGACAGACCGGG 576

RESULT 14
 AV616675 605 bp mRNA linear EST 28-NOV-2001
 LOCUS AV616675 Bos taurus ovary fetus Bos taurus cDNA clone E10V006G02
 DEFINITION 5', mRNA sequence.
 ACCESSION AV616675
 VERSION AV616675.1 GI:9752345
 KEYWORDS EST.
 SOURCE cow.
 ORGANISM Bos taurus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
 Bovidae; Bovinae; Bos.

REFERENCE 1 (bases 1 to 605)
 Takasuga, A., Hirotsune, S., Itoh, R., Jitohzono, A., Suzuki, H., Aso, H.
 and Sugimoto, Y.
 Establishment of a high throughput EST sequencing system using
 poly(A) tail-removed cDNA libraries and determination of 36,000
 bovine ESTs

JOURNAL Nucleic Acids Res. 29 (22), E108 (2001)
 MEDLINE 21570554
 COMMENT Contact: Yoshikazu Sugimoto
 Animal Genetics Division
 Shitakawa Institute of Animal Genetics
 Odatekura, Nishigo, Nishi-shitakawa, Fukushima 961-8061, Japan
 Tel: 81-248-25-5641
 Fax: 81-248-25-5725

Email: kazusugie@cocoa.ocn.ne.jp
 Single pass sequencing.
 This clone was obtained from a polyA-deleted cDNA library.

FEATURES
 source
 1..605
 Location/Qualifiers
 /organism="Bos taurus"
 /db_xref="taxon:9913"
 /clone_lib="E10V006G02"
 /clone_lib="Bos taurus ovary fetus"
 /tissue_type="ovary"
 /dev_stage="fetus"
 /lab_host="DH10B"
 /note="Vector: pZL1; Site_1: SalI; Site_2: NotI; Poly A
 was deleted from a NotI site"

BASE COUNT 158 a 128 c 148 g 171 t
 ORIGIN

Query Match 11.2%; Score 479; DB 10; Length 605;
 Best Local Similarity 88.9%; Pred. No. 7e-73;
 Matches 518; Conservative 0; Mismatches 65; Indels 0; Gaps 0;

QY 2556 CAGGCATTTATATCTTATGCTTATGCTGCAATTAACATTTTACCTTACCTAGCAATG 2615
 DB 10 CGGCGCATTTATATATCCATCTATAGCTGCGAGTTAAACATTTCTCTGTTAGCAATG 69
 QY 2616 TACCCATCATTTGCAATGAGAGAGTGTGAATGAAATGTTGCTGAGACAGCACTGA 2675
 DB 70 TACCCATCATTTGCAATGAGAGAGTGTGAATGAAATGTTGCTGAGACAGCACTGA 129
 QY 2676 AGATATAGAAAGAGCTAGAGAGCTGGAGAGATTGCTACAGAAAGCCATGCAAAACTTCC 2735
 DB 130 AGATATAGAAAGAGCTAGAGAGCTGGAGAGATTGCTACAGAAAGCCATGCAAAACTTCC 189

QY 2736 GAAGTGTGTTCTTTGACATCGGAGAGCAAGTTGAATACATGTATGCAAGAGTTTGC 2795
 DB 190 GGACTGTGTTCTTTTACTCGGAGAGAGAGTTTGAATATAGTATGCCAGAGTTTGC 249
 QY 2796 AAGTACCATACAGAAACTTTTGAGGAAAGCACACATCTTGGGGCTCATTTTGTATCA 2855
 DB 250 AGTATCCATACAGAAACTTTTGAGGAAAGCACACATCTTGGAAATTTACATTTGCCCTCA 309
 QY 2856 CCCAGGCAATGATGATTTTCTTCTATGCTGCTGTTCCGGTTGGTGCCATCTGTGTG 2915
 DB 310 CCCAGGCAATGATGATTTTCTTCTATGCTGCTGTTCCGGTTGGTGCCATCTGTGTG 369
 QY 2916 CAAATGATTCATGAACTTCATGATGTTGTTGTTGATCTCAGCATATGCTTGTG 2975
 DB 370 CACAAAGCATCATGAGATTTTCAAGATGTTCTCTTACTGTTCTCAGCTATGCTTGTG 429
 QY 2976 CCATGGCAGTGGGGCGAGTTCATTTGCTCCTGACTATGCCAAAGCAAGATGAG 3035
 DB 430 CCATGGCTGTGGGACAGTTCATTTGCTCCTGACTATGCCAAAGCAAGATGAG 489
 QY 3036 CAGCCACGTCATCATGATCATTTGAAGAAAGCCCTGATTCAGACGCTACGCCCTACG 3095
 DB 490 CAGCCACGTCATCATGATCATTTGAAGAAAGCCCTGATTCAGACGCTACGACAGAG 549
 QY 3096 GCCTCAGCCAAATACGTTTGAAGAAATGTGACATTTAATGA 3138
 DB 550 GCCTTAAGCCGAGTACAGTGAAGAAAGCGTGCAATTAATGA 592

RESULT 15
 BF313560 795 bp mRNA linear EST 21-NOV-2000
 LOCUS BF313560
 DEFINITION 601900192P1 NIH_MGC_19 Homo sapiens cDNA clone IMAGE:4129222 5',
 mRNA sequence.
 ACCESSION BF313560
 VERSION BF313560.1 GI:11261583
 KEYWORDS EST.
 SOURCE human.
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.

REFERENCE 1 (bases 1 to 795)
 NIH-MGC http://mgc.nci.nih.gov/.
 AUTHORS National Institutes of Health, Mammalian Gene Collection (MGC)
 TITLE Unpublished (1999)
 JOURNAL Contact: Robert Strausberg, Ph.D.
 COMMENT Email: c9apbs-r@mail.nih.gov
 Tissue Procurement: ATCC
 CDNA Library Preparation: Ling Hong/Rubin Laboratory
 CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNLN)
 DNA Sequencing by: Incyte Genomics, Inc.
 Clone distribution: MGC clone distribution information can be
 found through the I.M.A.G.E. Consortium/LNLN at: Image.lnl.gov
 Plate: L1CM1023 row: h column: 23
 High quality sequence stop: 674.

FEATURES
 source
 1..795
 Location/Qualifiers
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clone="IMAGE:4129222"
 /clone_lib="NIH_MGC_19"
 /tissue_type="neuroblastoma"
 /lab_host="DH10B (phage-resistant)"
 /note="Organ: brain; Vector: pORF; Site_1: XhoI; Site_2:
 EcoRI; CDNA made by oligo-dT priming. Directionally
 cloned into EcoRI/XhoI sites using the following 5'
 adaptor: GGCACGAG(G). Library constructed by Ling Hong
 in the laboratory of Gerald M. Rubin (University of
 California, Berkeley) using ZAP-CDNA synthesis kit
 (Stratagene) and Superscript II RT (Life Technologies).
 Note: this is a NIH_MGC library."

BASE COUNT 213 a 189 c 236 g 157 t

ORIGIN

Query Match	11.1%;	Score 475.2;	DB 12;	Length 795;
Best Local Similarity	83.6%;	Pred. No. 2.8e-72;		
Matches 612;	Conservative	0;	Mismatches 108;	Indels 12;
				Gaps 6;

QY	3065	AGCCCTCTGATTTGACAGCTTACAGCCCTTACGCGCTTCAAGCCCAATTAATGCTTGAAGAAT	3124
Db	2	AACCCCCTTATATACAGCTTACAGCAGGAAAGCCCTAATGCCAGAAC - ATGAAAGGAAT	60
QY	3125	GTGCATTTATGATGAGCGCTGTTCACTATCCACTGCACAGACATCCCGCTGCTCAG	3184
Db	61	GTCCATTTGGTGAAG--TGTATTCAACTATCCACCAGCCGACATCCAGTGTTCAG	118
QY	3185	GGCGTGAAGCTCGAAGTGAAGAAAGGCCAGACGCTGGCCCTCTGATAGGTATACATGGCTGT	3244
Db	119	GGACTGAGCCTGGAGGTGAAGAAAGGCCAGACGCGCTGTGTTGGGCAACAGTGGCTGT	178
QY	3245	GGGAAGAGCACAGTGTTCAGCTTCCCTTCAAGGCGCTTCATGACCCCTTGGCTGGTTCAGTG	3304
Db	179	GGGAAGAGCACAGTGTTCAGCTTCCCTTCAAGGCGCTTCATGACCCCTTGGGCAAGAAAGTG	238
QY	3305	CTAATGTATGTCGAAGAGATTAAGACACATCGAATGTCCAGTGGCTCCAGCACACTGGGC	3364
Db	239	CTGCTTGATGGCAAAAGAAATTAAGCGACATGATTTCAGTGGCTCCGAGCACCTGGGC	298
QY	3365	ATCTGTGTTCAGAGAGCCCATCTGTTTGACTGTGACATTTGCCGGAACATTTGCTATTGA	3424
Db	299	ATCTGTGTTCAGAGAGCCCATCTGTTTGACTGTGACATTTGCTGAGAACATTTGCTATTGA	358
QY	3425	GACAAACAGCCGGGTCTGTATCAATGAAAGATTTATCAGCAGCCAGCAAGAGGCCACATA	3484
Db	359	GACAAACAGCCGGGTGTCTACAGAGAGATGTGTGAGGACAGAAAGAGGCCACATA	418
QY	3485	CACACATTCATGAGACACTCCCTGAGAAATACAAACACAGATGTGAGACAAAGGAAC	3544
Db	419	CATGCTTCACTGATCTACTCTCTAATTAATATAGACATAAAGTGTGAGACAAAGAACT	478
QY	3545	CAGCTCTGTGTGCGCCAGAAACAGCGATTTGCCATAGCTCGCGCTTGTTTGACAGCT	3604
Db	479	CAGCTCTGTGTGCGCCAGAAACAGCGATTTGCCATAGCTCGTGTGCTTGTTTGACAGCT	538
QY	3605	CATATTTGCTTTTGTGATGAGAGCTACATCAGCTCTGTGATACAGAAAGTGAAGAAAGT--TGT	3663
Db	539	CATATTT--GCTTTGATTAAGCCACAGTCAAGCTCTGTGATACAGAAAGTGAAGAAAGTGTGCT	596
QY	3664	CCAAGAGAGCCCTGGAACAAAGCCAGAGAAAGCCGCACTGCATGTGTATAGGCCCAACCGCT	3723
Db	597	CCAAGAGAGCCCTGGAACAAAGCCAGAGAAAGGCGACCTCATGTATGTGTACCGC---T	652
QY	3724	GTCACCATTCAGAAATGCAAGATTAATGTGTGTTTCAGAAATGCGCAAGTCAAGGAGCA	3783
Db	653	GTCACCATTCAGAAATGCAAGATTAATGTGTGTTTCAGAAATGCGCAAGTCAAGGAGGC	710
QY	3784	TGGCAGACATCA 3795	
Db	711	ATGAGAGCATCA 722	

```
Search completed: December 10, 2002, 05:39:28
Job time : 29569 secs
```

OM protein - protein search, using sw model

Run on: December 9, 2002, 16:20:49 ; Search time 927 Seconds

(without alignments)
184.136 Million cell updates/sec

Title: US-09-672-725C-2
Perfect score: 6477

Sequence: 1 MDPEGGRKGSAEKNFWKMGK.....LLAQGIYFSMVSVQAGAKR 1281

Scoring table: BLOSUM62

Searched: 908470 seqs, 133250620 residues

Total number of hits satisfying chosen parameters: 908470

Minimum	DB	seq	length:	0
Maximum	DB	seq	length:	200000000000

Post-processing: Minimum Match 0%

Maximum	Match 100%
Listing first	45 summaries

Database : A_Geneseq_101002:*

1:	/SIDS1/gcgdata/genseq/genseq_emb1/AA1981.DAT *
2:	/SIDS1/gcgdata/genseq/genseq_emb1/AA1981.DAT *
3:	/SIDS1/gcgdata/genseq/genseq_emb1/AA1981.DAT *
4:	/SIDS1/gcgdata/genseq/genseq_emb1/AA1982.DAT *
5:	/SIDS1/gcgdata/genseq/genseq_emb1/AA1983.DAT *
6:	/SIDS1/gcgdata/genseq/genseq_emb1/AA1984.DAT *
7:	/SIDS1/gcgdata/genseq/genseq_emb1/AA1985.DAT *
8:	/SIDS1/gcgdata/genseq/genseq_emb1/AA1986.DAT *
9:	/SIDS1/gcgdata/genseq/genseq_emb1/AA1987.DAT *
10:	/SIDS1/gcgdata/genseq/genseq_emb1/AA1988.DAT *
11:	/SIDS1/gcgdata/genseq/genseq_emb1/AA1989.DAT *
12:	/SIDS1/gcgdata/genseq/genseq_emb1/AA1990.DAT *
13:	/SIDS1/gcgdata/genseq/genseq_emb1/AA1991.DAT *
14:	/SIDS1/gcgdata/genseq/genseq_emb1/AA1992.DAT *
15:	/SIDS1/gcgdata/genseq/genseq_emb1/AA1993.DAT *
16:	/SIDS1/gcgdata/genseq/genseq_emb1/AA1994.DAT *
17:	/SIDS1/gcgdata/genseq/genseq_emb1/AA1995.DAT *
18:	/SIDS1/gcgdata/genseq/genseq_emb1/AA1996.DAT *
19:	/SIDS1/gcgdata/genseq/genseq_emb1/AA1997.DAT *
20:	/SIDS1/gcgdata/genseq/genseq_emb1/AA1998.DAT *
21:	/SIDS1/gcgdata/genseq/genseq_emb1/AA1999.DAT *
22:	/SIDS1/gcgdata/genseq/genseq_emb1/AA2000.DAT *
23:	/SIDS1/gcgdata/genseq/genseq_emb1/AA2001.DAT *

Pred. NO. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	6477	100.0	1281	22	AAE00303	Dog (Pgp) P-glycoprotein
2	6469	99.9	1281	22	AAE00308	Dog P-glycoprotein
3	6463	99.8	1281	22	AAE00310	Dog P-glycoprotein
4	6456	99.7	1281	22	AAE00309	Dog P-glycoprotein
5	6420.5	99.1	1280	22	AAE00306	Dog P-glycoprotein
6	6420.5	99.1	1280	22	AAE00304	Dog P-glycoprotein
7	5861	90.5	1280	18	AAW44073	Human multidrug resistance protein 1
8	5861	90.5	1280	21	AAV58189	Human multidrug resistance protein 1
9	5861	90.5	1280	22	AAE01959	Human multidrug resistance protein 1
10	5861	90.5	1280	23	AAE19664	Human multidrug resistance protein 1

11	5861	90.5	1280	23	AAE07266	Human BCRP protein
12	5859	90.5	1280	23	AAE18967	Human P-gp mutant,
13	5857	90.4	1280	22	AAE181064	Cynomolgous monkey
14	5856	90.4	1280	23	AAE18968	Human P-gp mutant,
15	5854.5	90.4	1283	22	AAE18065	Cynomolgous monkey
16	5854	90.4	1280	23	AAE18966	Human P-gp mutant,
17	5853	90.4	1280	23	AAE18969	Human P-gp mutant,
18	5853	90.4	1280	23	AAE18970	Human P-gp mutant,
19	5853	90.4	1280	21	AAE18973	Human P-gp mutant,
20	5852	90.4	1280	21	AAE18972	Human P-gp mutant,
21	5852	90.4	1280	22	AAE18971	Human P-gp mutant,
22	5852	90.4	1280	22	AAE00306	Human P-glycoprotein
23	5852	90.4	1280	23	AAE24211	Human P-glycoprotein
24	5852	90.4	1280	23	AAE18965	Human MDR protein,
25	5852	90.4	1280	23	AAE07267	Human P-gp mutant,
26	5850	90.3	1280	23	AAE18977	Human BCRP protein
27	5849	90.3	1280	8	AAE07452	Human P-gp mutant,
28	5849	90.3	1280	23	AAE18971	Sequence encoded b
29	5847	90.3	1280	23	AAE18978	Human P-gp mutant,
30	5846	90.3	1280	11	AAE04868	Human P-gp mutant,
31	5845	90.2	1280	23	AAE18997	Protein encoded by
32	5844.5	90.2	1279	22	AAE11067	Human P-gp mutant,
33	5844.5	90.2	1279	22	AAE00307	Human P-glycoprotein
34	5844	90.2	1280	23	AAE18974	Human P-glycoprotein
35	5844	90.2	1280	23	AAE18979	Human P-gp mutant,
36	5844	90.2	1280	23	AAE18960	Human P-gp mutant,
37	5841	90.2	1280	23	AAE18972	Human P-gp mutant,
38	5840	90.2	1280	23	AAE18975	Human P-gp mutant,
39	5839	90.1	1280	15	AAE18974	Human P-gp mutant,
40	5836	90.1	1280	15	AAE18974	Human P-glycoprotein
41	5836	90.1	1280	22	AAE04347	Sequence encoded b
42	5835.5	90.1	1279	22	AAE00294	Human multidrug re
43	5832	90.0	1280	13	AAE18976	Mutated human P-g
44	5825	89.9	1280	23	AAE18975	Human P-gp mutant,
45	5825	89.9	1280	23	AAE04997	Wild-type human P
					Human P-glycoprotein	

ALIGNMENTS

Accession	Protein	Species	Accession	Protein	Species
AAE00303	standard; Protein; 1281 AA.		AAE00303	standard; Protein; 1281 AA.	
AAE00303			AAE00303		
13-JUN-2001	(first entry)		13-JUN-2001	(first entry)	
Dog (PCP)	P-glycoprotein (genotype C) #1.		Dog (PCP)	P-glycoprotein (genotype C) #1.	
Dog; P-glycoprotein; PCP; multidrug transporter; MDR1;			Dog; P-glycoprotein; PCP; multidrug transporter; MDR1;		
drug bioavailability; transgenic animal; genetic model.			drug bioavailability; transgenic animal; genetic model.		
Canis familiaris.			Canis familiaris.		
WO200123540-A2.			WO200123540-A2.		
05-APR-2001.			05-APR-2001.		
28-SEP-2000; 2000WO-US26767.			28-SEP-2000; 2000WO-US26767.		
28-SEP-1999; 99US-0156510.			28-SEP-1999; 99US-0156510.		
(GENT-) GENTEST CORP.			(GENT-) GENTEST CORP.		
Stocker PJ, Steinmetz DT, Crespi CL, Reif TC, Patten CJ;			Stocker PJ, Steinmetz DT, Crespi CL, Reif TC, Patten CJ;		
WPI; 2001-235373/24.			WPI; 2001-235373/24.		
N-PDB; AAD03488.			N-PDB; AAD03488.		
New dog P-glycoproteins (PCP) and their encoding nucleic acids, useful			New dog P-glycoproteins (PCP) and their encoding nucleic acids, useful		
for determining the bioavailability of drugs and for screening for dog			for determining the bioavailability of drugs and for screening for dog		
PCP inhibitors -			PCP inhibitors -		

XX Claim 17, Page 64-66; 11pp; English.
 CC The invention relates to dog P-glycoprotein (PGP) also referred
 CC as multidrug transporter (MDR1) and nucleic acids encoding them.
 CC The invention also includes fragments and biologically functional
 CC variants of dog P-glycoprotein. PGP and their nucleic acids are
 CC useful for determining the bioavailability of drugs and for
 CC screening PGP inhibitors. They are useful for the diagnosis and
 CC treatment of conditions characterised by PGP activity, by
 CC reducing or increasing PGP activity in a cell. PGP nucleic acids
 CC are used as oligonucleotide probes. Complements of PGP nucleic
 CC acids are used as antisense oligonucleotides, to induce a PGP
 CC 'knockout' phenotype. They are used to prepare a non-human
 CC transgenic animal, which are valuable as genetic models for
 CC human diseases.
 CC The present sequence is dog P-glycoprotein (PGP) also referred
 CC as genotype C protein. The PGP enzyme functions as an efflux
 CC pump exporting small molecules across the cell membrane. This
 CC enzyme is a member of the ABC transporter family.

XX Sequence 1281 AA:

Query Match 100.0%; Score 6477; DB 22; Length 1281;

Best Local Similarity 100.0%; Pred. No. 0; Mismatches 0; Indels 0; Gaps 0;

Matches 1281; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MDPEGGRGSAEKNFMWKKSKKNEKKKPTVSTFAMERYSNMLDRLMLVGTMAII 60
 DB 1 MDPEGGRGSAEKNFMWKKSKKNEKKKPTVSTFAMERYSNMLDRLMLVGTMAII 60
 QY 61 HGAALPLMLLVFGNMTDSFANAGISRNKTFPVYINSEINNNQHFNHLEEMTYVAYY 120
 DB 61 HGAALPLMLLVFGNMTDSFANAGISRNKTFPVYINSEINNNQHFNHLEEMTYVAYY 120
 QY 121 SGISAGVLAAYIOVSFWCLAGROILIKRQPFHAIMROEIGMFDVHDVDELNTRLTD 180
 DB 121 SGISAGVLAAYIOVSFWCLAGROILIKRQPFHAIMROEIGMFDVHDVDELNTRLTD 180
 QY 121 SGISAGVLAAYIOVSFWCLAGROILIKRQPFHAIMROEIGMFDVHDVDELNTRLTD 180
 DB 121 SGISAGVLAAYIOVSFWCLAGROILIKRQPFHAIMROEIGMFDVHDVDELNTRLTD 180
 QY 181 VSKINEGIGDKIGMFHSIATFTGTIVGFTRGMKLTLLVLAISPYVLGISAIAWAKILSS 240
 DB 181 VSKINEGIGDKIGMFHSIATFTGTIVGFTRGMKLTLLVLAISPYVLGISAIAWAKILSS 240
 QY 181 VSKINEGIGDKIGMFHSIATFTGTIVGFTRGMKLTLLVLAISPYVLGISAIAWAKILSS 240
 DB 181 VSKINEGIGDKIGMFHSIATFTGTIVGFTRGMKLTLLVLAISPYVLGISAIAWAKILSS 240
 QY 241 FTDKELAYAKAGAAVEEVLAAITPTVATAPGOKKELEERNKNEAKGIGIKKAITANIS 300
 DB 241 FTDKELAYAKAGAAVEEVLAAITPTVATAPGOKKELEERNKNEAKGIGIKKAITANIS 300
 QY 301 IGAAPLTIYASYALAFWGTISLVLSSEYSIGOVLTIVFESVLIGAFSIGASPSIEAFANA 360
 DB 301 IGAAPLTIYASYALAFWGTISLVLSSEYSIGOVLTIVFESVLIGAFSIGASPSIEAFANA 360
 QY 361 RGAAYEIRKIIDNKPISIDYSKSHKPDNIKGNLEFNKVFSPSRREVYILGLMLKYO 420
 DB 361 RGAAYEIRKIIDNKPISIDYSKSHKPDNIKGNLEFNKVFSPSRREVYILGLMLKYO 420
 QY 421 SGOYVALVNGSGCKSTVOLMORLYDPTDGMVCIDGQDRTINVRILREITGVSOEPEY 480
 DB 421 SGOYVALVNGSGCKSTVOLMORLYDPTDGMVCIDGQDRTINVRILREITGVSOEPEY 480
 QY 481 LEATTIAENIRYGRNVTMDTEIKAVKANAAYDFIMKLPNKFTLVBERAQISGGOKOR 540
 DB 481 LEATTIAENIRYGRNVTMDTEIKAVKANAAYDFIMKLPNKFTLVBERAQISGGOKOR 540
 QY 541 IATAALVANKPKILLDDATSAIDTESPAVVOVALDKARKGRTIVIAHRLSTVYRANAYI 600
 DB 541 IATAALVANKPKILLDDATSAIDTESPAVVOVALDKARKGRTIVIAHRLSTVYRANAYI 600
 QY 601 AAFDDGVITYEKNHDELKKEKGIYFKLVTMOTRGNETIELENATGESKSESDALEMSPKDS 660
 DB 601 AAFDDGVITYEKNHDELKKEKGIYFKLVTMOTRGNETIELENATGESKSESDALEMSPKDS 660
 QY 661 GSSLIKRSTRSIAHAPGQDRKLTGKEDLNNPVPVSFKILKLNSTWMPYFVVGIFCA 720
 DB 661 GSSLIKRSTRSIAHAPGQDRKLTGKEDLNNPVPVSFKILKLNSTWMPYFVVGIFCA 720

DB 661 GSSLIKRSTRSIAHAPGQDRKLTGKEDLNNPVPVSFKILKLNSTWMPYFVVGIFCA 720
 QY 721 IINGLQPAFSTIERSRIIGITFRQEDPETKRONSNMVSFLVIGIISFTFFLOGFTFG 780
 DB 721 IINGLQPAFSTIERSRIIGITFRQEDPETKRONSNMVSFLVIGIISFTFFLOGFTFG 780
 QY 781 KAGELITKRLRWVRSMRLRODVSFDDPKMTGALTTRLANDAAQVKAIGSLAVITQ 840
 DB 781 KAGELITKRLRWVRSMRLRODVSFDDPKMTGALTTRLANDAAQVKAIGSLAVITQ 840
 QY 841 NIANLGTGIIISLIGWOLTLILAIVPTIAYIVYEKMLSGALDKKELBAGKAT 900
 DB 841 NIANLGTGIIISLIGWOLTLILAIVPTIAYIVYEKMLSGALDKKELBAGKAT 900
 QY 901 EATENRTVASTLRQKFEYMYAOSLOVPYRNSLRKAHIFGVSPSITQAMMYFSYACFR 960
 DB 901 EATENRTVASTLRQKFEYMYAOSLOVPYRNSLRKAHIFGVSPSITQAMMYFSYACFR 960
 QY 961 FGAYLVANEFMNFODVLVFSALVFGAMAVGOVSSFADYAKAKVSAHVIMIEKSPLI 1020
 DB 961 FGAYLVANEFMNFODVLVFSALVFGAMAVGOVSSFADYAKAKVSAHVIMIEKSPLI 1020
 QY 1021 DSYSPHGLKPNITLQGNVTNEVEFNPPRPDIPLVQGLSLEVKKGOTLALVGSSGCKST 1080
 DB 1021 DSYSPHGLKPNITLQGNVTNEVEFNPPRPDIPLVQGLSLEVKKGOTLALVGSSGCKST 1080
 QY 1081 VVOLLERFYDPLAGSVLIDGKEIKHLNOMLRANLGIYSQEPILFDCSIAENIAYGNSR 1140
 DB 1081 VVOLLERFYDPLAGSVLIDGKEIKHLNOMLRANLGIYSQEPILFDCSIAENIAYGNSR 1140
 QY 1141 VVSHEEIMQAAKANIHFIETLPEKYNTRVGDKGQOLSGGOKORLATAALVROPHLL 1200
 DB 1141 VVSHEEIMQAAKANIHFIETLPEKYNTRVGDKGQOLSGGOKORLATAALVROPHLL 1200
 QY 1201 LDEKTSALDRESKRYOALDKAREGRTGIVIAHRLSTIONADLLIVPQNGKVKRSHGTHQ 1260
 DB 1201 LDEKTSALDRESKRYOALDKAREGRTGIVIAHRLSTIONADLLIVPQNGKVKRSHGTHQ 1260
 QY 1261 QLLAQKGIYFSMVSVOGAKR 1281
 DB 1261 QLLAQKGIYFSMVSVOGAKR 1281

RESULT 2
 AAE00308
 ID AAE00308 standard; protein; 1281 AA.
 AC AAE00308;
 DT 13-JUN-2001 (first entry)
 DE Dog P-glycoprotein (PGP) allelic variant (Genotype A).
 KW Dog; P-glycoprotein allelic variant; PGP; multidrug transporter;
 KW MDR1; drug bioavailability; transgenic animal; genetic model.
 OS Canis familiaris.
 XX Key location/Qualifiers
 XX Misc-difference 197 /note="His of genotypEC substituted by Gln"
 XX W0200123540-A2.
 XX 05-APR-2001.
 XX 28-SEP-2000; 2000WO-US26767.
 XX 28-SEP-1999; 99US-0156510.
 XX (GENT-) GENTEST CORP.
 XX Stocker PJ, Steimel-crespi DT, Crespi CL, Relif TC, Patten CJ;

XX WP1: 2001-235373/24.
DR N-PSDB: AAD03504.
XX
PT New dog P-glycoproteins (Pgp) and their encoding nucleic acids, useful
PT for determining the bioavailability of drugs and for screening for dog
PT Pgp inhibitors -
XX
PS Claim 17: Page 91-93; 11pp; English.
XX
CC The invention relates to dog P-glycoprotein (Pgp) also referred
CC as multidrug transporter (MDR1) and nucleic acids encoding them.
CC The invention also includes fragments and biologically functional
CC variants of dog P-glycoprotein. Pgp and their nucleic acids are
CC useful for determining the bioavailability of drugs and for
CC screening Pgp inhibitors. They are useful for the diagnosis and
CC treatment of conditions characterised by Pgp activity, by
CC reducing or increasing Pgp activity in a cell. Pgp nucleic acids
CC are used as oligonucleotide probes. Complements of Pgp nucleic
CC 'knockout' phenotype. They are used to prepare a non-human
CC transgenic animal, which are valuable as genetic models for
CC human diseases.
CC The present sequence is dog P-glycoprotein (Pgp) allelic variant.
CC This sequence is also referred as Genotype A protein. The Pgp
CC enzyme functions as an efflux pump exporting small molecules
CC across the cell membrane. This enzyme is a member of the ABC
CC transporter family.
XX
XX Sequence 1281 AA;
SQ
Query Match 99.9%; Score 6469; DB 22; Length 1281;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 1280; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 1 MDPEGGRGSAEKNFWMKKSKKNEKKKKPTVSTFAMFRYSNMDRLVMTGMAAT 60
DB 1 MDPEGGRGSAEKNFWMKKSKKNEKKKKPTVSTFAMFRYSNMDRLVMTGMAAT 60
QY 61 HGAALPLMLVFGNMTDFANAGISRNKTPVAINESITNTOHFINHEEMTAYYYY 120
DB 61 HGAALPLMLVFGNMTDFANAGISRNKTPVAINESITNTOHFINHEEMTAYYYY 120
QY 121 SGIGAGVLAAYIQVSWFCLAGROILKIRKOFHAIMROEIGMFVHYGELNRLTDD 180
DB 121 SGIGAGVLAAYIQVSWFCLAGROILKIRKOFHAIMROEIGMFVHYGELNRLTDD 180
QY 181 VSKINEGIGKIGMFHSHATFFTGFTVGRMKLTLVLAISPLVLSAIAIMAKTSS 240
DB 181 VSKINEGIGKIGMFHSHATFFTGFTVGRMKLTLVLAISPLVLSAIAIMAKTSS 240
QY 241 FTDKELIAYAKAGVAEEVLAIRTYIAFGQKKELERNKMLEAKGIGIKKAITANIS 300
DB 241 FTDKELIAYAKAGVAEEVLAIRTYIAFGQKKELERNKMLEAKGIGIKKAITANIS 300
QY 301 IGAFLITVASYALAFWYGTSLVLSSEYSIGOVLYFFSVLIGASISGASPTAEAFANA 360
DB 301 IGAFLITVASYALAFWYGTSLVLSSEYSIGOVLYFFSVLIGASISGASPTAEAFANA 360
QY 361 RGAAYEFKTIIDKPPSIDYSKSGHPRDINKNLEFKNVHESYPSRKREVKILKGLNLKVQ 420
DB 361 RGAAYEFKTIIDKPPSIDYSKSGHPRDINKNLEFKNVHESYPSRKREVKILKGLNLKVQ 420
QY 421 SGGTVALVNSGGKSTVQLMORLYDPDQWVCIDGODIRITINRHRETGYVSOEPV 480
DB 421 SGGTVALVNSGGKSTVQLMORLYDPDQWVCIDGODIRITINRHRETGYVSOEPV 480
QY 481 LRATTAENIRIRGRENVMDLEKAVKEANAVDTMLPNKEDTLVYERGAOLSGGOKOR 540
DB 481 LRATTAENIRIRGRENVMDLEKAVKEANAVDTMLPNKEDTLVYERGAOLSGGOKOR 540
QY 541 IAIARALVNRPKILLDEATSAIDTESEAVVOVALDKARKGRRTIVIAHRLSTVRNADVI 600
DB 541 IAIARALVNRPKILLDEATSAIDTESEAVVOVALDKARKGRRTIVIAHRLSTVRNADVI 600

DB 541 IAIARALVNRPKILLDEATSAIDTESEAVVOVALDKARKGRRTIVIAHRLSTVRNADVI 600
QY 601 AGFDDEVIYEKGNHDELMKEKGYFLVMTQMRGMEIELENAATGSKSESDALEMSPKDS 660
DB 601 AGFDDEVIYEKGNHDELMKEKGYFLVMTQMRGMEIELENAATGSKSESDALEMSPKDS 660
QY 661 GSSLIRKSTRNSIAHPQODRKLGTEDLNENVPVPSFWRIKLNSTEMPEYFVVGJCA 720
DB 661 GSSLIRKSTRNSIAHPQODRKLGTEDLNENVPVPSFWRIKLNSTEMPEYFVVGJCA 720
QY 721 IINGSLQPAFSIIFSRIGIIFTRDEDPEKRNQSMFVYLVLVIGIISFTIFFLQGFPG 780
DB 721 IINGSLQPAFSIIFSRIGIIFTRDEDPEKRNQSMFVYLVLVIGIISFTIFFLQGFPG 780
QY 781 KAGEILTRKRLRYMFRSMRLQDVSWFDDPKNTGALTRLNDAQVKAIGSLAVITQ 840
DB 781 KAGEILTRKRLRYMFRSMRLQDVSWFDDPKNTGALTRLNDAQVKAIGSLAVITQ 840
QY 841 NIANIGTGIIISLTYGMOLTLILAIVPIIAAGVENKMLSGQALDKKELEGAGKIAT 900
DB 841 NIANIGTGIIISLTYGMOLTLILAIVPIIAAGVENKMLSGQALDKKELEGAGKIAT 900
QY 901 EAIENFRVSLTRBQKFEYMYAOSLOVPYRNSLRKAHIFGVSPSTQAMMYTSYACFR 960
DB 901 EAIENFRVSLTRBQKFEYMYAOSLOVPYRNSLRKAHIFGVSPSTQAMMYTSYACFR 960
QY 961 FGAYLVANEFNFDVLLVFSALVFGAMAVGVSSFAPDVKAKVSAHVIIMEKSPIL 1020
DB 961 FGAYLVANEFNFDVLLVFSALVFGAMAVGVSSFAPDVKAKVSAHVIIMEKSPIL 1020
QY 1021 DSYSPHGKPTNLGNTFNFNVPYPRDIPVLOGLSLEVKKGOTLAVSSGCGKST 1080
DB 1021 DSYSPHGKPTNLGNTFNFNVPYPRDIPVLOGLSLEVKKGOTLAVSSGCGKST 1080
QY 1081 VOLLERFYDPLAGSVLIDGKEIKHLNWQMLRAHLGIVSEPIILFPCSTAEINAIYGDNR 1140
DB 1081 VOLLERFYDPLAGSVLIDGKEIKHLNWQMLRAHLGIVSEPIILFPCSTAEINAIYGDNR 1140
QY 1141 VVSHBEIMOAKKANIHHEFTLEPEKNTVSGDKGQLSGGOKORAIARALVROPHILL 1200
DB 1141 VVSHBEIMOAKKANIHHEFTLEPEKNTVSGDKGQLSGGOKORAIARALVROPHILL 1200
QY 1201 LDEATSAIDTESEKVVQVQALDKAREGRTCVIAHRLSTIONADLVVPONGVKEHGHQ 1260
DB 1201 LDEATSAIDTESEKVVQVQALDKAREGRTCVIAHRLSTIONADLVVPONGVKEHGHQ 1260
QY 1261 QLLAQKGIYSWVSQAGAKR 1281
DB 1261 QLLAQKGIYSWVSQAGAKR 1281
RESULT 3
AAE00309
ID AAE00309 standard; Protein: 1281 AA.
XX
AC AAE00309;
XX
DT 13-JUN-2001 (first entry)
XX
DE Dog P-glycoprotein (Pgp) allelic variant (Genotype B).
XX
KW Dog; P-glycoprotein allelic variant; Pgp; multidrug transporter;
XX MDR1; drug bioavailability; transgenic animal; genetic model.
OS Canis familiaris.
XX
FH Key Location/Qualifiers
FT Misc-difference 25
FT /note= "Asn of Genotypec substituted by Lys"
FT Misc-difference 197
FT /note= "His of Genotypec substituted by Gln"
XX
XX W0200123540-A2.

XX 05-APR-2001.
 PD 28-SEP-2000; 2000WO-US26767.
 XX 28-SEP-1999; 99US-0156510.
 XX (GENT-) GENEST CORP.
 PA Stocker PJ, Steimel-crespi DT, Crespi CL, Reif TC, Patten CJ;
 PI WPI; 2001-235373/24.
 DR N-PSDB; AAD03505.
 XX New dog P-glycoproteins (PGP) and their encoding nucleic acids, useful
 PT for determining the bioavailability of drugs and for screening for dog
 PT PGP inhibitors.
 PS Claim 17; Page 99-102; 111pp; English.
 XX The invention relates to dog P-glycoprotein (PGP) also referred
 CC as multidrug transporter (MDR1) and nucleic acids encoding them.
 CC The invention also includes fragments and biologically functional
 CC variants of dog P-glycoprotein. PGP and their nucleic acids are
 CC useful for determining the bioavailability of drugs and for
 CC screening PGP inhibitors. They are useful for the diagnosis and
 CC treatment of conditions characterised by PGP activity, by
 CC reducing or increasing PGP activity in a cell. PGP nucleic acids
 CC are used as oligonucleotide probes. Complements of PGP nucleic
 CC acids are useful as antisense oligonucleotides, to induce a PGP
 CC 'knockout' phenotype. They are used to prepare a non-human
 CC transgenic animal, which are valuable as genetic models for
 CC human diseases.
 CC The present sequence is dog P-glycoprotein (PGP) allelic variant.
 CC This sequence is also referred as Genotype B protein. The PGP
 CC enzyme functions as an efflux pump exporting small molecules
 CC across the cell membrane. This enzyme is a member of the ABC
 CC transporter family.
 XX Sequence 1281 AA;
 XX
 Query Match 99.8%; Score 6463; DB 22; Length 1281;
 Best Local Similarity 99.8%; Pred. No. 0;
 Matches 1279; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 421 SGQTVAVLNGSGGCKSTTVQLMQRLVDPDGMVCLDGODIRITINVRHLREITGVVSGQPV 480
 DB 421 SGQTVAVLNGSGGCKSTTVQLMQRLVDPDGMVCLDGODIRITINVRHLREITGVVSGQPV 480
 QY 481 LFATTIAENIRYGRNVTMDIEKAVKANAYDFIMKLPNKEDTLVGERGALSGGOKOR 540
 DB 481 LFATTIAENIRYGRNVTMDIEKAVKANAYDFIMKLPNKEDTLVGERGALSGGOKOR 540
 QY 541 IAIKALVRNPKIILDEATSAIDTESAVVOVALDKARKGRTIYIAHRLSTVRNADVI 600
 DB 541 IAIKALVRNPKIILDEATSAIDTESAVVOVALDKARKGRTIYIAHRLSTVRNADVI 600
 QY 601 AGPDDGVIVKGNHDELMKEKGIYFKLVMTQRTGNEIELEENATGESKSSDALENSPRDS 660
 DB 601 AGPDDGVIVKGNHDELMKEKGIYFKLVMTQRTGNEIELEENATGESKSSDALENSPRDS 660
 QY 661 GSSLIKRRSTRRSIHAPOGODRKIGTEDENENVPVPSFMRILKLKSTMPFVVGIFCA 720
 DB 661 GSSLIKRRSTRRSIHAPOGODRKIGTEDENENVPVPSFMRILKLKSTMPFVVGIFCA 720
 QY 721 IINGLOPAPSIIFRSIIIGIPTREDEPETRKONSMFSVLVLVGIISFTFFLOGFTEG 780
 DB 721 IINGLOPAPSIIFRSIIIGIPTREDEPETRKONSMFSVLVLVGIISFTFFLOGFTEG 780
 QY 781 KAGELITRRLRYMFRSMLRQDVSMFDDPKNTTCALITRIANDAAOYKKAIGSRLAVITQ 840
 DB 781 KAGELITRRLRYMFRSMLRQDVSMFDDPKNTTCALITRIANDAAOYKKAIGSRLAVITQ 840
 QY 841 NIANGTGIISLTYQMOLITLLAIYPIIAIGAVENKMLSGALKDKKELGACKIAT 900
 DB 841 NIANGTGIISLTYQMOLITLLAIYPIIAIGAVENKMLSGALKDKKELGACKIAT 900
 QY 901 EAIENFRVVSILTRQKEEYMAQSLQVYRNSLRKAHIGVSFTIQAMMYFSYAGCFR 960
 DB 901 EAIENFRVVSILTRQKEEYMAQSLQVYRNSLRKAHIGVSFTIQAMMYFSYAGCFR 960
 QY 961 FGAIYVANEFMFOVLLVESAIYVGAMAVGVSFADYAKAVSAHYIMITEKSPLI 1020
 DB 961 FGAIYVANEFMFOVLLVESAIYVGAMAVGVSFADYAKAVSAHYIMITEKSPLI 1020
 QY 1021 DSYSPHGLKPNLLEGNVTENEVFNYPTRPDIPLYLGSLSEYKGGQTLAVGSSGCKST 1080
 DB 1021 DSYSPHGLKPNLLEGNVTENEVFNYPTRPDIPLYLGSLSEYKGGQTLAVGSSGCKST 1080
 QY 1081 VVOLLERFYDPLAGSVLIDGKEIKHLNVQMLRAHLGIVSOEPIFDGSAENIAYGDSNR 1140
 DB 1081 VVOLLERFYDPLAGSVLIDGKEIKHLNVQMLRAHLGIVSOEPIFDGSAENIAYGDSNR 1140
 QY 1141 VVSHHEIMQAAKANIHHFLETLPEKYNTRVGDKGTOLSGGOKORIAIARALVRQPHIL 1200
 DB 1141 VVSHHEIMQAAKANIHHFLETLPEKYNTRVGDKGTOLSGGOKORIAIARALVRQPHIL 1200
 QY 1201 LDEATSAIDTESERVOEALDKARSGRTCIYIAHRLSTIONADLLIVFONKVKREHGHQ 1260
 DB 1201 LDEATSAIDTESERVOEALDKARSGRTCIYIAHRLSTIONADLLIVFONKVKREHGHQ 1260
 QY 1261 QLLAQKGIYFSMVSVOAGAKR 1281
 DB 1261 QLLAQKGIYFSMVSVOAGAKR 1281

RESULT 4
 ID AAE00310
 ID AAE00310 standard; Protein: 1281 AA.
 AC AAE00310;
 AC AAE00310;
 DT 13-JUN-2001 (first entry)
 XX Dog P-glycoprotein (PGP) allelic variant (Genotype D).
 XX Dog; P-glycoprotein allelic variant; PGP; multidrug transporter;
 KW MDR1; drug bioavailability; transgenic animal; genetic model.
 KW

Query Match	Best local Similarity	Matches 1277:	Conservative	2:	Mismatches	2:	Indels	0:	Gaps
1	MDPEGGRKSAEKAKFWKMKRKKSKKNEKKKPEYTFAMFYSNNWLDRLYMLVGTMAII	60							
2	MDPEGGRKSAEKAKFWKMKRKKSKKNEKKKPEYTFAMFYSNNWLDRLYMLVGTMAII	60							
3	HGAALPILMLVFGWMTDSFANAISGNTKFPYIIINESTTNNOHFINHLEEMTTAYYY	120							
4	HGAALPILMLVFGWMTDSFANAISGNTKFPYIIINESTTNNOHFINHLEEMTTAYYY	120							
5	SGIGAGVLYAAVIOYSWFCLAGRQLIKRQFPHALIMROEISGFVVDVGEINRLTDD	180							
6	SGIGAGVLYAAVIOYSWFCLAGRQLIKRQFPHALIMROEISGFVVDVGEINRLTDD	180							
7	VSKINSGIDKIGMFHSIATFEFTGVIQFTRGKNTLVLTAISPLYGSAINAKTISS	240							

Db	181	VSKINSGDKTGMFFQSIATFFPFVGTBKGKMLTVILAIISPVLGSLAAWKLSS	240
Qy	241	FTDKELLATAKGAAYAEVLAIRTVIAPGGOKKELERYKNLBEAKGICIKAKITAMIS	300
Db	241	FTDKELLATAKGAAYAEVLAIRTVIAPGGOKKELERYKNLBEAKGICIKAKITAMIS	300
Qy	301	IGAFAFLIYAASVALFMWGTSLVJSESEYJGQULVFEFSGVLIGAFSSJGQASPSIEAFANA	360
Db	301	IGAFAFLIYAASVALFMWGTSLVJSESEYJGQULVFEFSGVLIGAFSSJGQASPSIEAFANA	360
Qy	361	RGAAEYELFKIDNKSIDSYSGSKHCKEDNKGYLEKRNHVSYSRRREVILKGLMLKYQ	420
Db	361	RGAAEYELFKIDNKSIDSYSGSKHCKEDNKGYLEKRNHVSYSRRREVILKGLMLKYQ	420
Qy	421	SGQYVALVNGSGCKSTVVOJLMQRLXPDPDGMCTIDGODIRTVIHRHLREITGVSOEYV	480
Db	421	SGQYVALVNGSGCKSTVVOJLMQRLXPDPDGMCTIDGODIRTVIHRHLREITGVSOEYV	480
Qy	481	LFATTAIENIRYGRBNVTMDEIERAAYKANAYDFIMKLEPNKPDVLGBERGAOLSGGOKOR	540
Db	481	LFATTAIENIRYGRBNVTMDEIERAAYKANAYDFIMKLEPNKPDVLGBERGAOLSGGOKOR	540
Qy	541	IATARALVRNPKLILDEBATSALDTESEAYVOVALDKAKGRTTVIAHRSLSTRANADYI	600
Db	541	IATARALVRNPKLILDEBATSALDTESEAYVOVALDKAKGRTTVIAHRSLSTRANADYI	600
Qy	601	AGPDDGYIEKNGHDELMEKESGYIEFKLVYMQTNGNIELENNATGSKESSEPALEMSPKOS	660
Db	601	AGPDDGYIEKNGHDELMEKESGYIEFKLVYMQTNGNIELENNATGSKESSEPALEMSPKOS	660
Qy	661	GSSILIKRRSTRSIIAHPOGOKRKLGTIEDENENVPVSWRLKLNSTEMPVYVIFCA	720
Db	661	GSSILIKRRSTRSIIAHPOGOKRKLGTIEDENENVPVSWRLKLNSTEMPVYVIFCA	720
Qy	721	IINGLOPAPSIIESRILIGITFRDEDEBETKRONSMFVYLVLGILISITFTFLGSGFTFG	780
Db	721	IINGLOPAPSIIESRILIGITFRDEDEBETKRONSMFVYLVLGILISITFTFLGSGFTFG	780
Qy	781	KAGEILTKRLRIVNVRFSMLRODVSMFDPDPKNTTGALTTLRLANDAQAOKYAGISRLAVITQ	840
Db	781	KAGEILTKRLRIVNVRFSMLRODVSMFDPDPKNTTGALTTLRLANDAQAOKYAGISRLAVITQ	840
Qy	841	NIANLGGIILISITGMQTLTLLLAUYPIIATAGVYEMKMLSGQALDKDKKEJGACKINT	900
Db	841	NIANLGGIILISITGMQTLTLLLAUYPIIATAGVYEMKMLSGQALDKDKKEJGACKINT	900
Qy	901	EALENPRTVYSLTPROKFEWYAAOSLOVPARNSLBRANLHFVSSISIOAMMYSVAGCR	960
Db	901	EALENPRTVYSLTPROKFEWYAAOSLOVPARNSLBRANLHFVSSISIOAMMYSVAGCR	960
Qy	961	FGAYLVANEPNFODVLLVSAIYFEGAMAVGOVSSAPDIYAKAVSAAHYIMITEKSPIL	1020
Db	961	FGAYLVANEPNFODVLLVSAIYFEGAMAVGOVSSAPDIYAKAVSAAHYIMITEKSPIL	1020
Qy	1021	DYSVPHGLKPNLLEGNVTENFENVENYPTRPDIPLVLOGLSLEVKKGQTLALVSSSGCKST	1080
Db	1021	DYSVPHGLKPNLLEGNVTENFENVENYPTRPDIPLVLOGLSLEVKKGQTLALVSSSGCKST	1080
Qy	1081	VVOLLEFRDPLAGSVLLDGEIKELHLNWQRLRAHLAIGYSOEPIILFDCSISENIAIYADNDR	1140
Db	1081	VVOLLEFRDPLAGSVLLDGEIKELHLNWQRLRAHLAIGYSOEPIILFDCSISENIAIYADNDR	1140
Qy	1141	VVSHREELMOAKENIMHIFETLPEKNTPVGSKGLOLSSGOKORIMATARALROPHILL	1200
Db	1141	VVSHREELMOAKENIMHIFETLPEKNTPVGSKGLOLSSGOKORIMATARALROPHILL	1200
Qy	1201	LDEATSLADTESEKYVOALDKAKBEGTCIVIAHRSLSTIONADLIYVFOGKYEKGHTHQ	1260
Db	1201	LDEATSLADTESEKYVOALDKAKBEGTCIVIAHRSLSTIONADLIYVFOGKYEKGHTHQ	1260
Qy	1261	QLLAQKGIYFSMVSVOAGAKR 1281	
Db	1261	QLLAQKGIYFSMVSVOAGAKR 1281	


```

RESULT 6
AAE00304
ID AAE00304 standard; Protein; 1280 AA.
XX
AC AAE00304;
XX
XX 13-JUN-2001 (first entry)
XX
XX Dog P-glycoprotein (PGP) #2.
XX
XX Dog: P-glycoprotein; PGP; multidrug transporter; MDR1;
XX drug bioavailability; transgenic animal; genetic model.
OS
XX Canis familiaris.
XX
XX WO200123540-A2.
XX
XX 05-APR-2001.
XX
XX 28-SEP-2000; 2000WO-0526767.
XX
XX 28-SEP-1999; 99US-0156510.
XX
XX (GENT-) GENTEST CORP.
XX
XX Stocker PJ, Steimel-crespi DT, Crespi CL, Reif TC, Patten CJ;
XX WPI; 2001-235373/24.
XX
XX N-PSDB; AAD03489.
XX
XX New dog P-glycoproteins (PGP) and their encoding nucleic acids, useful
XX for determining the bioavailability of drugs and for screening for dog
XX PGP inhibitors.
XX
XX Example 2; Page 72-75; 11pp; English.
XX
XX The invention relates to dog P-glycoprotein (PGP) also referred
XX as multidrug transporter (MDR1) and nucleic acids encoding them.
XX The invention also includes fragments and biologically functional
XX variants of dog P-glycoprotein. PGP and their nucleic acids are
XX useful for determining the bioavailability of drugs and for
XX screening PGP inhibitors. They are useful for the diagnosis and
XX treatment of conditions characterised by PGP activity, by
XX reducing or increasing PGP activity in a cell. PGP nucleic acids
XX are used as oligonucleotide probes. Complements of PGP nucleic
XX 'knockout' phenotype. They are used to prepare a non-human
XX transgenic animal, which are valuable as genetic models for
XX human diseases.
XX
XX The present sequence is dog P-glycoprotein (PGP). The
XX PGP enzyme functions as an efflux pump exporting small molecules
XX across the cell membrane. This enzyme is a member of the ABC
XX transporter family.
XX
XX Sequence 1280 AA:
XX
Query Match 99.1%; Score 6420.5; DB 22; Length 1280;
Best Local Similarity 99.3%; Pred. No. 0;
Matches 1272; Conservative 4; Mismatches 4; Indels 1; Gaps 1;
QY 1 MDEPGGKGSAAEKNFWMKSKSKKNEKKKPTSTFMFRYSNNLDRLYLVGTMAII 60
DB 1 MDEPGGKGSAAEKNFWMKSKSKK-EKKEKPTSTFMFRYSNNLDRLYLVGTMAII 59
QY 61 HGAALPLMLVGNMTDSFANNGISRNKTPVYIINESTITNTQHFIIHLEEMTTAYYY 120
DB 60 HGAALPLMLVGNMTDSFANNGISRNKTPVYIINESTITNTQHFIIHLEEMTTAYYY 119
QY 121 SGIGAGVLAAYIYVSFMCIAAGROIILKIRKOFHAIIMROEIGFVDVHVGELNTRLTDD 180
DB 120 SGIGAGVLAAYIYVSFMCIAAGROIILKIRKOFHAIIMROEIGFVDVHVGELNTRLTDD 179

```

```

QY 181 VSKINEGIDKIGMFHSIATFTGFIYGTGRMKTLLIYLAISPYLGLSAAIMAKILSS 240
DB 180 VSKINEGIDKIGMFHSIATFTGFIYGTGRMKTLLIYLAISPYLGLSAAIMAKILSS 239
QY 241 FTDEKELAYAKAGAAVEEYLAIRVIAGGOKKELERYNKNLEBAKIGIKKAITANIS 300
DB 240 FTDEKELAYAKAGAAVEEYLAIRVIAGGOKKELERYNKNLEBAKIGIKKAITANIS 299
QY 301 IGAAPFLIYASVALAFWYGSILVSEYSIGOVLPVFPSYLIAGFSIQOASPIEAPANA 360
DB 300 IGAAPFLIYASVALAFWYGSILVSEYSIGOVLPVFPSYLIAGFSIQOASPIEAPANA 359
QY 361 RGAAYEIFKTIIDNKPSIDSYSGSKHPDNKGNLEFFKNVHFSYSPSRKREYIKGLNKKYQ 420
DB 360 RGAAYEIFKTIIDNKPSIDSYSGSKHPDNKGNLEFFKNVHFSYSPSRKREYIKGLNKKYQ 419
QY 421 SGOYVALVGNSSGCKSTYVLMORLYDPTDGMVCIQODIRITVNRHLREITGVSGQEPY 480
DB 420 SGOYVALVGNSSGCKSTYVLMORLYDPTDGMVCIQODIRITVNRHLREITGVSGQEPY 479
QY 481 LEATTIAENIRGRBNVTMDIEKAKEANAAYDEIMKLPMKPTLIVGERGAOLSGOQXOR 540
DB 480 LEATTIAENIRGRBNVTMDIEKAKEANAAYDEIMKLPMKPTLIVGERGAOLSGOQXOR 539
QY 541 IAIARALVRNPKIILLDEATSAIDTSEAVVOVALDKARKGRTTIVIAHRLSTVRNADVI 600
DB 540 IAIARALVRNPKIILLDEATSAIDTSEAVVOVALDKARKGRTTIVIAHRLSTVRNADVI 599
QY 601 AGFDGVIVKGNHDELMKEKGIYFKLVYMTQTRNGNEIELENAATESKSDALEMSPDS 660
DB 600 AGFDGVIVKGNHDELMKEKGIYFKLVYMTQTRNGNEIELENAATESKSDALEMSPDS 659
QY 661 GSSLIKRRSTRSHAPOGDRKLGTEKEDLENVBSVSWRLIKLNSTEMPEYVVGIFCA 720
DB 660 GSSLIKRRSTRSHAPOGDRKLGTEKEDLENVBSVSWRLIKLNSTEMPEYVVGIFCA 719
QY 721 IINGLOPASISIFSRIGIFTRDEDEPETKRONSNMFVYLVGIIISIFIFFILOGTFG 780
DB 720 IINGLOPASISIFSRIGIFTRDEDEPETKRONSNMFVYLVGIIISIFIFFILOGTFG 779
QY 781 KAGEIILTKRLRYVFRSMRLRODVSNFDDPKNTGALTTRLANDAAOVKAGIATVITQ 840
DB 780 KAGEIILTKRLRYVFRSMRLRODVSNFDDPKNTGALTTRLANDAAOVKAGIATVITQ 839
QY 841 NINMLGTGIIISLYGWLTLALLAIVPIIAAGVVEKMLSGOALKDKKELBAGKAT 900
DB 840 NINMLGTGIIISLYGWLTLALLAIVPIIAAGVVEKMLSGOALKDKKELBAGKAT 899
QY 901 EAIENFRVVSILTRQKFEYMAAOSLOVPRNSLRKAHIGVSFSITQAMMYPSVAGCFR 960
DB 900 EAIENFRVVSILTRQKFEYMAAOSLOVPRNSLRKAHIGVSFSITQAMMYPSVAGCFR 959
QY 961 FGAYIVANFEMFQDVLVFSAIYFGAMAVGOVSSFADPAKAKVSAAHYIMITESSPI 1020
DB 960 FGAYIVANFEMFQDVLVFSAIYFGAMAVGOVSSFADPAKAKVSAAHYIMITESSPI 1019
QY 1021 DSYSPHGLKPTLIEGNTYFENVVNPTRPDIVYDGLSLEPAKGGTALVSSGCGKST 1080
DB 1020 DSYSPHGLKPTLIEGNTYFENVVNPTRPDIVYDGLSLEPAKGGTALVSSGCGKST 1079
QY 1081 VVOLLERYDPLAGSVLIDKEIKHLNVOMLRAHLVISOEPLTFPCSAIENIAGVDSNR 1140
DB 1080 VVOLLERYDPLAGSVLIDKEIKHLNVOMLRAHLVISOEPLTFPCSAIENIAGVDSNR 1139
QY 1141 VVSHEEIMQAKEANIHHFJETLPEKYNTRYDGKGTQLSGGOKORIAIARALVRPHILL 1200
DB 1140 VVSHEEIMQAKEANIHHFJETLPEKYNTRYDGKGTQLSGGOKORIAIARALVRPHILL 1199
QY 1201 LDEATSAIDTESKVVQOALDKAREGRCTIVIAHRLSTIQNADILIVYONKVEHEHTQ 1260
DB 1200 LDEATSAIDTESKVVQOALDKAREGRCTIVIAHRLSTIQNADILIVYONKVEHEHTQ 1259
QY 1261 QLLAKGIIYFSMVSVOGAKR 1281

```

DB 1260 QLLAQKGYFSMISVQAGAKR 1280

RESULT 7

AA04073 standard; protein; 1280 AA.

AA04073:

26-JUN-1998 (first entry)

Human multidrug resistance P-glycoprotein MDRI.

Human; multidrug resistance P-glycoprotein; MDRI; prokaryotic homologue.

Homo sapiens.

Key Location/Qualifiers

Region 1..640 "MDRI-N from Fig 1"

Region 641..1280 "MDRI-C from Fig 1"

W09740160-A1.

30-OCT-1997.

24-APR-1997; 97MO-N000216.

24-APR-1996; 96EP-0201094.

(UKR-) RIJXSUNIV GRONINGEN.

Bolhuis H, Konings WN, Van Veen HW, Venema K;

WPI; 1997-535844/49.

Prokaryotic homologue of human multiple drug resistance protein -

used to screen for compounds that inhibit, or avoid, drug resistance

Claim 10: Fig 1; 35pp: English.

The present invention describes a recombinant or isolated nucleic acid (1), derived from a prokaryotic gene, which encodes at least a specific and/or functional part of a transporter protein (TP), or its derivatives, which has functional and/or structural similarity with the P-glycoprotein (Pg) encoded by the human multidrug resistance (MDR)1 gene. The present sequence represents the human MDRI protein, derived from MDRI-N and MDRI-C as shown in the specification in figure 1. (1) is used to express recombinant proteins; its fragments are also useful as probes and primers for detection and amplification of related DNA. The protein produced, or cells expressing them, are used to determine if substances can inhibit, or avoid, MDR proteins, and in a screening method for identifying compounds that inhibit transport of cytotoxic substances from cells. Also, cells with a transmembrane protein, especially where expressed from (1), can provide additional MDR, particularly for use as a model system to study mechanisms of action of Pg.

Sequence 1280 AA:

Query Match 90.5%; Score 5861; DB 18; Length 1280; Best Local Similarity 90.7%; Pred. No. 0; Matches 1163; Conservative 55; Mismatches 60; Indels 4; Gaps 4;

DB 1 MDPGSGRGS-A-EKNFMKSKSKNEKKKPYSTFAFRYSNNLDRMYLGMMAI 59
1 MDLSDNRNGAKKRNKFNKSKER-DKREKPYVSFSMRYSDKLIMVAVGLAI 59
60 IRGAALPLMLLVFGNMTDSFANAGISRNKTEPVIIINSTNTQHTINLLEEMTYAVY 119
60 IHGAGLPLMLLVFGEMTIDIFANAG-NLDELMSNITNRSNDINDIGFTFM-TLEDMRYAVY 117

QY 120 YSGIGAGLVAAAYIOVSFWCLAAAGROILKIRKQFFHAIIMROEIGMDVDHDELNTRLTD 179
DB 118 YSGIGAGLVAAAYIOVSFWCLAAAGROILKIRKQFFHAIIMROEIGMDVDHDELNTRLTD 177
QY 180 DVSKINGIGDKIGMFPHSIAFTFTGFTVGTFRGMKLTIVILASIPVLGISAIAINAKILIS 239
DB 178 DVSKINGIGDKIGMFPHSIAFTFTGFTVGTFRGMKLTIVILASIPVLGISAIAINAKILIS 237
QY 240 SFTDKELLATAKAGAAVEEYLAIRTVIAFGGOKKELEKRNKULBEAKIGIKAKITANI 299
DB 238 SFTDKELLATAKAGAAVEEYLAIRTVIAFGGOKKELEKRNKULBEAKIGIKAKITANI 297
QY 300 SIGAFLILYASALAFMYGTSLVLSSEYSIGOVLVFFSVLIGAFSIGOASPSIEAFAN 359
DB 298 SIGAFLILYASALAFMYGTSLVLSSEYSIGOVLVFFSVLIGAFSIGOASPSIEAFAN 357
QY 360 ARGAAVEIFKIIDNKPISIDYSKSGHKPDNIKGLKFNKVNHFSPSRKBEVILKLNKLV 419
DB 358 ARGAAVEIFKIIDNKPISIDYSKSGHKPDNIKGLKFNKVNHFSPSRKBEVILKLNKLV 417
QY 420 QSGOTVALVNSCGKSTTVOLMORLYDPTDGMVCLDGOIRIINRHLREITGVYSGEP 479
DB 418 QSGOTVALVNSCGKSTTVOLMORLYDPTDGMVCLDGOIRIINRHLREITGVYSGEP 477
QY 480 VLFATTTAENIRYGRBNVTMDIEKAVKEANAAYDFIMKLBNKEDTLVGBRGAQLSGGOKQ 539
DB 478 VLFATTTAENIRYGRBNVTMDIEKAVKEANAAYDFIMKLBNKEDTLVGBRGAQLSGGOKQ 537
QY 540 RIATAARLVNPNPKIILLDEATSAIDTESEAVVOVALDKARKGRTTVIAHRLSTYRNDV 599
DB 538 RIATAARLVNPNPKIILLDEATSAIDTESEAVVOVALDKARKGRTTVIAHRLSTYRNDV 597
QY 600 IAGFDGQIVYEKNHDELKKEGITEFKLVYMQTNGNEIELEANA TGSKSESALMSPKD 659
DB 598 IAGFDGQIVYEKNHDELKKEGITEFKLVYMQTNGNEIELEANA TGSKSESALMSPKD 657
QY 660 SGSSLIKRSTRSRISIAPOGODRKLTGREDENENVPVPSFMRILKLNSTEMPYPVVGIFC 719
DB 658 SRSSLIKRSTRSRISIAPOGODRKLTGREDENENVPVPSFMRILKLNSTEMPYPVVGIFC 717
QY 720 AINGGLOPAFSLIFSRIGITFTRDEPETKRONSNFSLFVLGIIISFTFFLOGTFE 779
DB 718 AINGGLOPAFSLIFSRIGITFTRDEPETKRONSNFSLFVLGIIISFTFFLOGTFE 777
QY 760 GKAGEIILTKRLRYMVRSMLRDQVSWPDDPKNTGALTTRFLANDAQVGAIGSLAVIT 839
DB 778 GKAGEIILTKRLRYMVRSMLRDQVSWPDDPKNTGALTTRFLANDAQVGAIGSLAVIT 837
QY 840 ONIANLGTGIIISLIGMOLTLTLLAIVPIIAIGAVVEMKLSGQALKDKKELBGAGKIA 899
DB 838 ONIANLGTGIIISLIGMOLTLTLLAIVPIIAIGAVVEMKLSGQALKDKKELBGAGKIA 897
QY 900 TEALENRTVVSILTRQKREYVYAOISLOVPRNSLRKANHFGVSFTTOMMYFYACAF 959
DB 898 TEALENRTVVSILTRQKREYVYAOISLOVPRNSLRKANHFGVSFTTOMMYFYACAF 957
QY 960 RFQAYIVANEFMFOVLLVYESAIVFGMAVGOVSPAPDIYAKAKSAHAHYMIETKSP 1019
DB 958 RFQAYIVANEFMFOVLLVYESAIVFGMAVGOVSPAPDIYAKAKSAHAHYMIETKSP 1017
QY 1020 IDSYPHGLKPNLLEGNVTFNEVFNPTRPDIPLVLOGLSLEVKKGQTLALVSSGCGKS 1079
DB 1018 IDSYPHGLKPNLLEGNVTFNEVFNPTRPDIPLVLOGLSLEVKKGQTLALVSSGCGKS 1077
QY 1080 TVVOLLERFYPDLAGSVLLDGEIKRLNVOQLRAHGLIYQOEPIIPDGSIAENIAYGNS 1139
DB 1078 TVVOLLERFYPDLAGSVLLDGEIKRLNVOQLRAHGLIYQOEPIIPDGSIAENIAYGNS 1137
QY 1140 RVYSHEEIMQAKAENIHFETPEKYNTRVGDGKTQLSGGOKRIATARALVROPHIL 1199
DB 1138 RVYSHEEIMQAKAENIHFETPEKYNTRVGDGKTQLSGGOKRIATARALVROPHIL 1197

QY 1200 LIDEATSAIDTESKVVQDQALDKAREGRCTIVIAHRLSTIONADLIYFONGKYKEHGT 1259
 Db 1198 LIDEATSAIDTESKVVQDQALDKAREGRCTIVIAHRLSTIONADLIYFONGKYKEHGT 1257
 OY 1260 QQLAOKGIYFSWVSVOGAKR 1281
 Db 1258 QQLAOKGIYFSWVSVOGAKTR 1279

RESULT 8
 AAY58186
 ID AAY58186 standard; Protein: 1280 AA.
 AC AAY58186;
 XX
 XX
 DT 14-MAR-2000 (first entry)
 DE Human wild-type multidrug resistance-1 (MDR-1) protein.
 XX
 KM Multidrug resistance; MDR-1; P-glycoprotein;
 KM Transmembrane efflux pump; haematopoietic stem cell; transduction;
 KM bone marrow transplantation; chemotherapy; radiation therapy; cancer;
 KM gene therapy; gene replacement; genetic defect; thalassaemia;
 KM Gaucher's disease; sickle cell anaemia; leukaemia; ex vivo expansion;
 KM cytokine; wild-type.
 XX
 OS Homo sapiens.
 XX
 FH Key Location/Qualifiers
 FT Misc-difference 185
 FT /note- "this residue is val in a mutant MDR-1
 FT (AAY58187)"
 FT
 XX W09961589-A2.
 XX
 PD 02-DEC-1999.
 XX
 XX 27-MAY-1999; 99WO-US11825.
 PF
 XX 28-MAY-1998; 98US-0086988.
 PR
 XX (SUJ-) ST JUDE CHILDREN'S RES HOSPITAL.
 PA
 PI Sorrentino B, Bunting K;
 XX
 XX WPI, 2000-072615/06.
 DR N-PSDB; AA249332.
 DR
 XX Ex vivo expansion of haematopoietic stem cells transduced with a
 PT sequence encoding human multidrug resistance-1, used for bone marrow
 PT transplantation -
 XX
 PS Claim 10; Page 71-79; 113pp; English.
 XX
 CC This sequence represents human wild-type multidrug
 CC resistance protein MDR-1. MDR-1 is a transmembrane efflux
 CC pump, responsible for the export of drugs from certain
 CC cells, particularly cancer cells. Wild-type MDR-1 shows
 CC increased resistance to etoposide and decreased resistance to vinca
 CC alkaloids compared with a mutant form (AAY58187) where the Gly at
 CC position 185 is replaced by Val. The invention relates to transducing
 CC haematopoietic stem cells with nucleic acid encoding an MDR protein
 CC and culturing the modified cells. The modified haematopoietic stem
 CC cells are useful in bone marrow transplantation (to reconstitute
 CC haematopoietic systems in patients who have undergone chemotherapy or
 CC radiation therapy) and in ex vivo gene therapy of genetic defects in
 CC cells derived from haematopoietic stem cells, e.g., thalassaemia,
 CC Gaucher's disease, sickle cell anaemia or leukaemia. The modified
 CC cells can also be used to identify factors involved in regulating
 CC proliferation and differentiation in haematopoietic stem cells.
 CC Haematopoietic stem cells that express MDR-1 will be protected against
 CC chemotherapeutic agents, so can be engrafted while the patient is
 CC undergoing chemotherapy. Expansion of (rare) haematopoietic stem cells

CC provides sufficient cells to permit standard biochemical analysis.
 CC Overexpression of MDR-1 allows cytokine-driven expansion of
 CC haematopoietic stem cells by at least 10-fold compared with a maximum
 CC of 4-fold in known procedures.
 CC
 XX Sequence 1280 AA;
 SO

Query Match 90.5%; Score 5861; DB 21; Length 1280;
 Best Local Similarity 90.7%; Pred. No. 0;
 Matches 1163; Conservative 55; Mismatches 60; Indels 4; Gaps 4;

QY 1 MDPEGRKGSAA-EKEMKMGCKSKKNEKKEKPPVSTFAFRRYSNMUJLMLVGMAT 59
 Db 1 MDLEDGRNGAKKKNFKFNKSEK-DKEREKPTVSFMSRRYSNMUJLMLVGMAT 59
 OY 60 IHGAALPLMLLVFGNMTDPSFANAGISRNKTPPVITINSEITNNQHFNIHEEMTAYAY 119
 Db 60 IHGAGLPLMLLVFGNMTDPSFANAG-NLEDLMSNTNNSDINDGFFRN-LEEDMTAYAY 117
 OY 120 YSGIAGVLYVAAYIOVSFWCLAAGRQILKIRKQFHAIMROEIGWPDVHDVGLNTRLD 179
 Db 118 YSGIAGVLYVAAYIOVSFWCLAAGRQILKIRKQFHAIMROEIGWPDVHDVGLNTRLD 177
 OY 180 DVSKINEGIGDKIGMFPSIATPFTGFIYVGTGPKMTLITIAISPYLGSAAIMAKIIS 239
 Db 178 DVSKINEGIGDKIGMFPSIATPFTGFIYVGTGPKMTLITIAISPYLGSAAIMAKIIS 237
 OY 240 SPTDKELLAAVAKAGAAVEVLAIRTVIAFGOKKELERYNKMLEBAKGIGIKATANI 299
 Db 238 SPTDKELLAAVAKAGAAVEVLAIRTVIAFGOKKELERYNKMLEBAKGIGIKATANI 297
 OY 300 SIGAFLITIAASYALAWYGTSLVLSSEYSIGOVLIYFVSVLIGAFSIQASPSIEAFAN 359
 Db 298 SIGAFLITIAASYALAWYGTSLVLSSEYSIGOVLIYFVSVLIGAFSVQASPSIEAFAN 357
 OY 360 ARGAAVEIFKIIDNKPISIDSYSKSGHKPDNIKGNLEFFKNVHSPSRKRVKILKGNLKV 419
 Db 358 ARGAAVEIFKIIDNKPISIDSYSKSGHKPDNIKGNLEFFKNVHSPSRKRVKILKGNLKV 417
 OY 420 QSGQVALVGNSSGCKSTVOLMQRLYDPTDGMVCIDGQDITRTINVRLREITGVVSQEP 479
 Db 418 QSGQVALVGNSSGCKSTVOLMQRLYDPTDGMVSVSDQDIRTINVRLREITGVVSQEP 477
 OY 480 VLPATTIAENIRYGREVNVMDIEKAVKENADFTMKLPKPEPTLVCGRQALSQGQK 539
 Db 478 VLPATTIAENIRYGREVNVMDIEKAVKENADFTMKLPKPEPTLVCGRQALSQGQK 537
 OY 540 RIARALVNRPKILLIDEATSAIDTESAAYVOVALDKARKGRTTIVIAHRLSTVRNADV 599
 Db 538 RIARALVNRPKILLIDEATSAIDTESAAYVOVALDKARKGRTTIVIAHRLSTVRNADV 597
 OY 600 IAFPDGVIYEKNDHMKKEGIYFKLVYMTQRTGNEIELENANGESKSESDALEMSKPD 659
 Db 598 IAFPDGVIYEKNDHMKKEGIYFKLVYMTQRTGNEIELENANGESKSESDALEMSKPD 657
 OY 660 SSSLLKRRSTRSIIAPOGODRKLTGEDLENNPVSFRRILKLNSTEMPYVVGIFC 719
 Db 658 SSSLLKRRSTRSIVGSAQDRKLTSTKALDESIPVSPWIKMLNTJEMPYFVVGIFC 717
 OY 720 AIIINGLQPAFSTIIFSRIGITRDEDETKRONNMKSVFLVYIGISFTFFLOGTF 779
 Db 718 AIIINGLQPAFALIFSKIIIGVTRIDDEPKRONNLSLFLALGIIISFTFFLOGTF 777
 OY 780 GRAGELITRLRYNFRSMLRODVSWPDDPKRTTALTTRLANDAAVYKAGISRLAVIT 839
 Db 778 GRAGELITRLRYNFRSMLRODVSWPDDPKRTTALTTRLANDAAVYKAGISRLAVIT 837
 OY 840 QNTANLGTGIIISLYGQTLTLLAIVPIAIAGVYEMKMLSGOALKDKKELGAGAKIA 899
 Db 838 QNTANLGTGIIISLYGQTLTLLAIVPIAIAGVYEMKMLSGOALKDKKELGAGAKIA 897
 OY 900 TTAIEFRIVYSLTEQKFEYMTAOSLOVYRNSLRANHIFGVSFITQAMTFYSYAGCP 959
 Db 900 TTAIEFRIVYSLTEQKFEYMTAOSLOVYRNSLRANHIFGVSFITQAMTFYSYAGCP 959

Db 898 TEAIEFRTVVSLETOEQEHEHMYAQLQVPRNSLRKAHIFGITEFSTQAMTFYAGCF 957
QY 960 REGAYLVANEFNNFODVLIVSAIVGAMAVGOVSRPAPDYAKKSAVAHVIIMIEKSP 1019
Db 958 REGAYLVANHLKMSFEDVLIVSAIVGAMAVGOVSRPAPDYAKKSAVAHVIIMIEKSP 1017
QY 1020 IDSYPHGLKPNLTGNTFENVEVFNYPTRPDIPVLOGLSLEVKKGOTLALVSSGGCKS 1079
Db 1018 IDSISTEGMLPNTLEGNNTFEGVFNYPTRPDIPVLOGLSLEVKKGOTLALVSSGGCKS 1077
QY 1080 TVVOLLEREPDPLAGSVLIDGKEIKHLNVOMLRALHGLIVSOEPIIFDCSIAENIAYGDS 1139
Db 1078 TVVOLLEREPDPLAGSVLIDGKEIKHLNVOMLRALHGLIVSOEPIIFDCSIAENIAYGDS 1137
QY 1140 RVYSHEIWOAKENIHFIETLEPEKYNTRYGDKGTOLSGOKORIALAVALROPHIL 1199
Db 1138 RVYSOHEIVRAKANIHAFTLESLPKYISTKVGDKGTOLSGOKORIALAVALROPHIL 1197
QY 1200 LIDEATSALDTESEKVVQBALDKAREGRTCIYIAHRLSTIONADLIVFONGKYEKGTH 1259
Db 1198 LIDEATSALDTESEKVVQBALDKAREGRTCIYIAHRLSTIONADLIVFONGKYEKGTH 1257
QY 1260 QQLLAOKGIYFSMVYQAGAKR 1281
Db 1258 QQLLAOKGIYFSMVYQAGAKR 1279

RESULT 9
AAB81959
ID AAB81959 standard; protein; 1280 AA.
XX
XX AAB81959;
XX
XX 02-JUL-2001 (first entry)
XX
XX Human MDRL.
XX
XX Human; MDRL; multi-drug resistance; cholesterol uptake;
XX
XX hypercholesterolemia; hypocholesterolemia; atherosclerosis;
XX
XX coronary artery disease; cerebral vascular disease.
XX
XX Homo sapiens.
XX
XX OS
XX
XX WO200121762-A2.
XX
XX PD
XX 29-MAR-2001.
XX
XX PF
XX 22-SEP-2000; 2000MO-US26099.
XX
XX PR
XX 23-SEP-1999; 99US-0155819.
XX
XX PA
XX (BARN-) BARNES-JEWISH HOSPITAL.
XX
XX PI
XX Stenson WF, Tessner T;
XX
XX WP1: 2001-328100/34.
XX
XX DR
XX Modulating cellular cholesterol uptake for treating hyper
XX
XX PT
XX hypocholesterolemia, involves administering an agent that inhibits or
XX
XX PT
XX increases the expression of multiple drug resistance-1
XX
XX PS
XX Example 1: Page 43-47; 47pp; English.
XX
XX CC
XX The present invention describes a method of modulating cholesterol uptake
XX
XX CC
XX in cells, involving administering an agent capable of inhibiting the
XX
XX CC
XX multi-drug resistance protein MDRL. This is useful in the prevention and
XX
XX CC
XX treatment of cholesterol-related diseases, including
XX
XX CC
XX hypercholesterolemia, hypocholesterolemia, atherosclerosis, coronary
XX
XX CC
XX artery disease and cerebral vascular diseases. The present sequence is
XX
XX CC
XX the human MDRL protein.
XX
XX SQ
XX Sequence 1280 AA;

Query Match 90.58; Score 5861; DB 22; Length 1280;
Best Local Similarity 90.78; Pred. No. 0;
Matches 1163; Conservative 55; Mismatches 60; Indels 4; Gaps 4;
QY 1 MDPEGRGSA-EKFMFMGKKSKKNEKEKPPVSTFAMFRYSNMIDRLYLVTMAAI 59
Db 1 MDLEGRGSAKKKFFFLNKKSEK-DKEKKPPVSVFMSFRYSNMIDRLYLVTMAAI 59
QY 60 IHGAALPLMLVFGMNTDSFANAGISRNKTPPVINSSTNNNOHFENHLEEMTYAY 119
Db 60 IHGAGPLMLVFGEMTDIFANAG-NEDLMSNTNNSDINDGFEFNK-LEEDHTIRAY 117
QY 120 YSGIGAGVVAAYIOVSFMCIAAGROILKIRKOFHAIMROEIGMPDVHVGELNRLTD 179
Db 118 YSGIGAGVVAAYIOVSFMCIAAGROILKIRKOFHAIMROEIGMPDVHVGELNRLTD 177
QY 180 DYSKINEGIDKIGMFHSIATFETGTVGTGKMLTVLTAISPVLGSAAMAKILS 239
Db 178 DYSKINEGIDKIGMFHSIATFETGTVGTGKMLTVLTAISPVLGSAAMAKILS 237
QY 240 SFTDKELAYAKAGAAVEVLAARTVTAFGOKKELERKNLEAKGIGIKKAITANI 299
Db 238 SFTDKELAYAKAGAAVEVLAARTVTAFGOKKELERKNLEAKGIGIKKAITANI 297
QY 300 SIGAFLILYASVALAFWYGTSLVLSSEYSGOVLTVFVSYLIGAFSIGQSPSIEAFAN 359
Db 298 SIGAFLILYASVALAFWYGTSLVLSSEYSGOVLTVFVSYLIGAFSIGQSPSIEAFAN 357
QY 360 ARGAAVEIFKIIDNKPISIDYSKSGHKPDNIKGNLEFNHVFSPSRREYVILGKLVK 419
Db 358 ARGAAVEIFKIIDNKPISIDYSKSGHKPDNIKGNLEFNHVFSPSRREYVILGKLVK 417
QY 420 QSGQTVALVNGSCGKSTTVQMLQRLYPTDGMVICIDODIRTIIVRLREITGVSOEP 479
Db 418 QSGQTVALVNGSCGKSTTVQMLQRLYPTDGMVICIDODIRTIIVRLREITGVSOEP 477
QY 480 VLPATTIAENIRYGRNVMTDEIEKAVKANAYDTIMKIPHEFDVLGERGAQSLGGKQ 539
Db 478 VLPATTIAENIRYGRNVMTDEIEKAVKANAYDTIMKIPHEFDVLGERGAQSLGGKQ 537
QY 540 RIATRALVNRPKILLDEATSALDTESEAVVQVALDKARKGRTIYIAHRLSTVRNADV 599
Db 538 RIATRALVNRPKILLDEATSALDTESEAVVQVALDKARKGRTIYIAHRLSTVRNADV 597
QY 600 IAGFDGVIVEKGNHDELMKEKGIYFKLVYTWOTRGNEIELENAAGESSESALDEMSKD 659
Db 598 IAGFDGVIVEKGNHDELMKEKGIYFKLVYTWOTRGNEIELENAAGESSESALDEMSKD 657
QY 660 SGSSLIRKSTRSIIAPQODRKLTGEDLNNVPVPSFWIRILKINSTEMPEYVAGIFC 719
Db 658 SRSLSIRKSTRSIIAPQODRKLTGEDLNNVPVPSFWIRILKINSTEMPEYVAGIFC 717
QY 720 AIIINGLOPAPFIIFFRRTIIGTRDEDPETKRONSNMFSVLVLVIGIISFTFVLQGTFF 779
Db 718 AIIINGLOPAPFIIFFRRTIIGTRDEDPETKRONSNMFSVLVLVIGIISFTFVLQGTFF 777
QY 780 GKAGEILTKRLRYWFRSMRLQOVSWFDDPKNTGALTTRLANDAAQVGAIGSLAVYT 839
Db 778 GKAGEILTKRLRYWFRSMRLQOVSWFDDPKNTGALTTRLANDAAQVGAIGSLAVYT 837
QY 840 ONINNLGIGIISLYGMOLTLTLALVPIITAIAGVEMKMSLSCGALKDKKELEGAKIA 899
Db 838 ONINNLGIGIISLYGMOLTLTLALVPIITAIAGVEMKMSLSCGALKDKKELEGAKIA 897
QY 900 TEAIEFRTVVSLETOEQEHEHMYAQLQVPRNSLRKAHIFGVSEITQAMTFYAGCF 959
Db 898 TEAIEFRTVVSLETOEQEHEHMYAQLQVPRNSLRKAHIFGVSEITQAMTFYAGCF 957
QY 960 REGAYLVANEFNNFODVLIVSAIVGAMAVGOVSRPAPDYAKKSAVAHVIIMIEKSP 1019
Db 958 REGAYLVANHLKMSFEDVLIVSAIVGAMAVGOVSRPAPDYAKKSAVAHVIIMIEKSP 1017
QY 1020 IDSYPHGLKPNLTGNTFENVEVFNYPTRPDIPVLOGLSLEVKKGOTLALVSSGGCKS 1079

|||||
DB 1018 IDSTSTEGSLMPVTLLEGNNVFEVNPYPTRPDIPLVQLSLSEVKKGGCLALVYSSGGKGS 1077
QY 1080 TVVOLLERFYDPLAGSVLIDGKEIKHLNVOMLRALHGLVISOEPILFPOCSIAENIAYDND 1139
DB 1078 TVVOLLERFYDPLAGSVLIDGKEIKRLNVOMLRALHGLVISOEPILFPOCSIAENIAYDND 1137
QY 1140 RVVSHBEIWOAKKANIHHEITLPEKYNTRYGDKGTQLSGGQKORIAIARALYRPHIL 1199
DB 1138 RVVSOEIVRAKKAENIHAFIESLPNKYSTKVGDKGTQLSGGQKORIAIARALYRPHIL 1197
QY 1200 LIDETSALDPESEKVVQVADLKAREGRTGCIYIAHRLSTIQNADLIYVFGNGVKEHGT 1259
DB 1198 LIDETSALDPESEKVVQVADLKAREGRTGCIYIAHRLSTIQNADLIYVFGNGVKEHGT 1257
QY 1260 QOLLAQKGIYFSMVSVQAGAKR 1281
DB 1258 QOLLAQKGIYFSMVSVQAGAKR 1279
RESULT 10
AAE18964
ID AAE18964 standard; Protein; 1280 AA.
AC AAE18964;
XX
DT 21-MAY-2002 (first entry)
XX
DE Human wild type P-glycoprotein (wt P-gp).
XX
KW Human: P-glycoprotein; vaccine; tumour; chemotherapeutic agent; cancer;
XX haematopoietic cell; cytotoxicity; cytosolic; P-gp.
XX Homo sapiens.
OS
XX MO200210205-A2.
PN 07-FEB-2002.
XX
PF 01-AUG-2001; 2001WO-US24560.
XX
PR 01-AUG-2000; 2000US-222313P.
XX
PA (UNIT) UNIV ILLINOIS FOUND.
PI Ruth A, Roninson I;
XX
DR WPI; 2002-206182/26.
XX
PT New human P-glycoprotein mutants for treating cancer, comprises an
PT ability to confer increased resistance to chemotherapeutic drugs
PT relative to wild type P-glycoprotein or P-glycoprotein having a glycine
PT to valine substitution
XX
PS Claim 44; Fig 3; 47pp; English.
XX
XX The present invention relates to an isolated mutant human P-glycoprotein
XX molecule. The invention is used as vaccine. An antibody specific for the
XX mutant P-glycoprotein is useful for determining whether a human tumour
XX sample comprises tumour cells expressing mutant human P-glycoprotein. A
XX recombinant expression construct comprising nucleic acid encoding mutant
XX P-glycoprotein is useful for administering an increased amount of a
XX chemotherapeutic agent to an individual with cancer, by transducing ex
XX vivo haematopoietic cells from the individual with cancer, with the recombinant
XX expression construct, reintroducing the transduced haematopoietic cells
XX into the individual and administering an increased amount of a
XX chemotherapeutic agent without consequent haematopoietic cytotoxicity.
XX The mutant P-glycoprotein is useful for treating and diagnosing cancer
XX and for screening compounds for the capacity to disrupt binding of
XX cytotoxic drugs with the mutant P-glycoprotein or to disrupt drug efflux
XX from cells expressing the mutant P-glycoprotein. The present sequence is
XX human wild type P-glycoprotein (wt P-gp).

Seq Sequence 1280 AA:
Query Match 90.5%; Score 5861; DB 23; Length 1280;
Best Local Similarity 90.7%; Pred. No. 0;
Matches 1163; Conservative 55; Mismatches 60; Indels 4; Gaps 4;
QY 1 MDPEGRGKSA-EKNFMKMGKSKKNEKKEKPPVSTFAFRRYSNMJDLRYMLVGTMAAT 59
DB 1 MDPEGRGKSA-EKNFMKMGKSKKNEKKEKPPVSTFAFRRYSNMJDLRYMLVGTMAAT 59
QY 60 IHGAALPLMLVFGNMTDSFANAGISRKKTPPVYINSSINNOHFNHLEEMTYTAY 119
DB 60 IHGAALPLMLVFGNMTDSFANAGISRKKTPPVYINSSINNOHFNHLEEMTYTAY 117
QY 120 YSGIGAGVLAAYIQVSWFCLAAGRQILKIRKOPFAHIMROEIGMPDVHVGELNRLTD 179
DB 118 YSGIGAGVLAAYIQVSWFCLAAGRQILKIRKOPFAHIMROEIGMPDVHVGELNRLTD 177
QY 180 DYSKINEGIGDKIGMFPSTATFEFTGFTVGTTRGKTLTYIALISPYLGSAAIMAKIIS 239
DB 178 DYSKINEGIGDKIGMFPSTATFEFTGFTVGTTRGKTLTYIALISPYLGSAAIMAKIIS 237
QY 240 SFTDKELLAYAKAGAAVEVLAAIRTVIAFGOKKELERYNKMLEAKGIGIKKAITANI 299
DB 238 SFTDKELLAYAKAGAAVEVLAAIRTVIAFGOKKELERYNKMLEAKGIGIKKAITANI 297
QY 300 SIGAFLILYASALAWYGTSLVSEYSIGOVLYVFSVLIGAFSIGQASPSIAPAN 359
DB 298 SIGAFLILYASALAWYGTSLVSEYSIGOVLYVFSVLIGAFSIGQASPSIAPAN 357
QY 360 ARGAAVEIFKIIDNKPISIDYSKSGHKPDNIKGNLEKFNWHEPSRKVKYILKGLNLY 419
DB 358 ARGAAVEIFKIIDNKPISIDYSKSGHKPDNIKGNLEKFNWHEPSRKVKYILKGLNLY 417
QY 420 QSGQTVLVGNSGCGKSTVQLMQRLYDPTDGNVYCIDGODIRTNVRLHREITGVSOEP 479
DB 418 QSGQTVLVGNSGCGKSTVQLMQRLYDPTDGNVYCIDGODIRTNVRLHREITGVSOEP 477
QY 480 VLFAITTAENIRYGRNVTYDDEIEKAVKENAYDFIMKLPKFFDVLVGRGAQLSGGKQ 539
DB 478 VLFAITTAENIRYGRNVTYDDEIEKAVKENAYDFIMKLPKFFDVLVGRGAQLSGGKQ 537
QY 540 RIATARALVNRPKILLDEATSLDPESEAVVOVALDKRKRRTTVIAHRLSTVRNDV 599
DB 538 RIATARALVNRPKILLDEATSLDPESEAVVOVALDKRKRRTTVIAHRLSTVRNDV 597
QY 600 IAGFDGVIYEKGNHDELKKEGIVFKLVYTMQTRNETIELENATGESKSESDALEMSPKD 659
DB 598 IAGFDGVIYEKGNHDELKKEGIVFKLVYTMQTRNETIELENATGESKSESDALEMSPKD 657
QY 660 SGSSILTKRSTRSIAPOGODRKLTGREDLWENVPVPSFRIILKLNSTEMPVYVIGTC 719
DB 658 SGSSILTKRSTRSIAPOGODRKLTGREDLWENVPVPSFRIILKLNSTEMPVYVIGTC 717
QY 720 AIINGGLOPAFSLIFSRIGITFRDDEPTEKRONSMFSLVLYGITSFTFFLOGTTF 779
DB 718 AIINGGLOPAFSLIFSRIGITFRDDEPTEKRONSMFSLVLYGITSFTFFLOGTTF 777
QY 780 GRAGELLTRKRYMRYSMLRODVSWFDDPKTGTALTRLANDAAOVYKAGISRLAVIT 839
DB 778 GRAGELLTRKRYMRYSMLRODVSWFDDPKTGTALTRLANDAAOVYKAGISRLAVIT 837
QY 840 ONIANLGTGIIISLYGMOULTLLLAIVPIIAAGVEMKMLSGQALKDKKELGACKIA 899
DB 838 ONIANLGTGIIISLYGMOULTLLLAIVPIIAAGVEMKMLSGQALKDKKELGACKIA 897
QY 900 TEAIENFTVSVLTFEDQKFEYVYAOGLVPRNSLRKAHIFGVFSITQAMMYSYAGCF 959
DB 898 TEAIENFTVSVLTFEDQKFEYVYAOGLVPRNSLRKAHIFGVFSITQAMMYSYAGCF 957
QY 960 RFGATVYANEFNPNDDVLLVPSATYFGAMAYGVSSFPADYAKKVAHAHYIMITEKSP 1019
DB 958 RFGATVYANEFNPNDDVLLVPSATYFGAMAYGVSSFPADYAKKVAHAHYIMITEKSP 1017

QY 1020 IDSYPHGIKPTLLEGNTEFNEVFNPTRPDIPVLOGLSLEVKGGOTLALVSSGGCKS 1079
DB 1018 IDSYSTEGMLPTLLEGNTEFNEVFNPTRPDIPVLOGLSLEVKGGOTLALVSSGGCKS 1077
QY 1080 IYVOLLERFYDPLAGSVLIDGKEIKHLNVOMLRKHLGIVSOEPIILFDCSIAENIAYGDS 1139
DB 1078 IYVOLLERFYDPLAGSVLIDGKEIKHLNVOMLRKHLGIVSOEPIILFDCSIAENIAYGDS 1137
QY 1140 RYVSHBEIQAQKANEINHHFETLPEKYNTRVGDKGQOLSGGOKRIALRAIYRPHIL 1199
DB 1138 RYVSOEEIYRAKANEINHHFETLPEKYNTRVGDKGQOLSGGOKRIALRAIYRPHIL 1197
QY 1200 LIDEATSALDTESEKVOEALDKAREGRTCIYIAHRLSTIONADLIYVFGNGKYEKGTH 1259
DB 1198 LIDEATSALDTESEKVOEALDKAREGRTCIYIAHRLSTIONADLIYVFGNGKYEKGTH 1257
QY 1260 QOLLAQGIYFSMVSVQAGAKR 1281
DB 1258 QOLLAQGIYFSMVSVQAGAKR 1279

RESULT 11
ABB07266
ID ABB07266 standard; Protein; 1280 AA.
XX
AC ABB07266;
XX
DT 26-MAR-2002 (first entry)
XX
DE Human BCRP protein related seq Id No. 2.
XX
KM Stem cell; ATP transport protein; ATP-binding cassette; antiparkinsonian;
KM hepatocytic; neurodegenerative; cytosolic; antianemic; muscular; BCRP;
KM cardiant; gene therapy.
XX
OS Homo sapiens.
XX
XX WO200192877-A2.
XX
PD 06-DEC-2001.
XX
PF 30-MAY-2001; 2001MO-US17459.
XX
PR 31-MAY-2000; 2000US-0584586.
PR 29-MAY-2001; 2001US-0866866.
XX
PA (SJUD-) ST JUDE CHILDREN'S RES HOSPITAL.
XX
PI Sorrentino B, Schuetz J;
XX
DR WPI: 2002-114368/15.
DR N-PSDB: ABA94365.
XX
PT Identifying a stem cell, for treating e.g., muscular dystrophy,
PT myocardial infarction, parkinson's disease, or neurodegenerative
PT disorders, comprises detecting the expression of an ATP transport
PT protein (BCRP) by a cell
XX
PS Disclosure; Page 55-58; 87pp; English.
XX
XX The invention provides a method of identifying and/or isolating a stem
XX cell that involves detecting the expression of an ATP transport protein
XX containing a conserved ATP-binding cassette (BCRP) by a cell in a sample
XX comprising stem cells. The isolated stem cells may be used in the
XX treatment of diseases such as muscular dystrophy, degenerative liver
XX disorder, myocardial infarction, parkinson's disease, degenerative
XX disorders of the brain, and for tissue regeneration or replacement.
XX Haematopoietic cells can be used in bone marrow transplants (e.g., for
XX treatment of leukemia) and for ex vivo gene therapy for treating blood
XX diseases such as sickle cell anemia and thalassemia. The stem cells can
XX also be used as cell targets in gene therapy protocols. The present
XX sequence represents a sequence related to the BCRP for which no relevant

CC information has been provided in the specification.
XX
SQ Sequence 1280 AA;
Query Match 90.5%; Score 5861; DB 23; Length 1280;
Best Local Similarity 90.7%; Pred. No. 0;
Matches 1163; Conservative 55; Mismatches 60; Indels 4; Gaps 4;
QY 1 MDPEGKRGSA-EKNFMKMGKSKKNEKKEKPPVSTFAMFRSNNDRLYMLVGMAT 59
DB 1 MDLEEDRMGAKKKNFNFKLNKSEK-DKEKKPPVSVFSMRISNNLMDKLIVAVGILAAI 59
QY 60 IHCALPLMLVFGMTQSPFANAGISRNKTPPVILINSITNTOHFTNLHEEETVAY 119
DB 60 IHCAGLPLMLVFGMTQSPFANAG-NLEDLMSNTINSNDINDTEFAN-LEEDTFRAY 117
QY 120 YSGIGAGVLAAYIQVSWCLAGROILKIRKOFPHAIMROETWEPVHYGELNTRLTD 179
DB 118 YSGIGAGVLAAYIQVSWCLAGROILKIRKOFPHAIMROETWEPVHYGELNTRLTD 177
QY 180 DYSKINIEGIDKIGMFFHSIAFEFTGFTVGRGKTLVLAISPVLGSAAIMAKILS 239
DB 178 DYSKINIEGIDKIGMFFHSIAFEFTGFTVGRGKTLVLAISPVLGSAAIMAKILS 237
QY 240 SFTDKELAYAKAGAAVEEVLAIKRTVIAFGQKKLEBRYNKNLEBAKIGICAKITANI 299
DB 238 SFTDKELAYAKAGAAVEEVLAIKRTVIAFGQKKLEBRYNKNLEBAKIGICAKITANI 297
QY 300 SIGAFLLIYASTALAFWYGISLVLSSEYSGQVITVFESVLIGAFSIGQSPSIEAFAN 359
DB 298 SIGAFLLIYASTALAFWYGISLVLSSEYSGQVITVFESVLIGAFSIGQSPSIEAFAN 357
QY 360 ARGAAYEIRKIIDNKPSIDISYSGSKHPDNKGNLEFNHVSYSRREKVALKGLNLKY 419
DB 358 ARGAAYEIRKIIDNKPSIDISYSGSKHPDNKGNLEFNHVSYSRREKVALKGLNLKY 417
QY 420 QSGQFVALVGNSSCGSKSTVQMLQRLYDPTDGMVCIDGODITITVNRILREITGVSE 479
DB 418 QSGQFVALVGNSSCGSKSTVQMLQRLYDPTDGMVCIDGODITITVNRILREITGVSE 477
QY 480 VLEATTIAENIRYGRENVIMDELEKAVKANAYDITMKLPNKFDTLVGERGAQSLSGQ 539
DB 478 VLEATTIAENIRYGRENVIMDELEKAVKANAYDITMKLPNKFDTLVGERGAQSLSGQ 537
QY 540 RIATARALVNPRIKILIDEATSALDTESEAVVQVALDKARKGRTTIVIAHRLSTVRNDV 599
DB 538 RIATARALVNPRIKILIDEATSALDTESEAVVQVALDKARKGRTTIVIAHRLSTVRNDV 597
QY 600 IAGFDGVIYKGNHDELMKEKGIYKLYTMOTRGNEIELENADESKSESDALEMSPKD 659
DB 598 IAGFDGVIYKGNHDELMKEKGIYKLYTMOTRGNEIELENADESKSESDALEMSPKD 657
QY 660 SGSSLIKRSGTRSRSHAPOGQDRKLTKEJDLNENVPVFSWIRILKINSTEMPYFVVGIFC 719
DB 658 SGSSLIKRSGTRSRSHAPOGQDRKLTKEJDLNENVPVFSWIRILKINSTEMPYFVVGIFC 717
QY 720 AIIINGGLQPAFSIIFSRKIIIFTRDEDPETKRONSMESVYLIVIGIISFTITFQGTFF 779
DB 718 AIIINGGLQPAFSIIFSRKIIIFTRDEDPETKRONSMESVYLIVIGIISFTITFQGTFF 777
QY 780 GKAEIILTKRLRYVWFSMSLRQDVSWPDDPKNTGALTTRRIANDAAQVGAIGSRILAVIT 839
DB 778 GKAEIILTKRLRYVWFSMSLRQDVSWPDDPKNTGALTTRRIANDAAQVGAIGSRILAVIT 837
QY 840 ONINLNGIIGIISITVEMQTLTLLALVPIITAIAGVEMKMLSGQALKDKKELEGAGKIA 899
DB 838 ONINLNGIIGIISITVEMQTLTLLALVPIITAIAGVEMKMLSGQALKDKKELEGAGKIA 897
QY 900 TEATENFTVSLTRKQEPFYMATQSTQVPYRNSLRKAHIFGVSEITQAMMYFSTAGCF 959
DB 898 TEATENFTVSLTRKQEPFYMATQSTQVPYRNSLRKAHIFGVSEITQAMMYFSTAGCF 957
QY 960 RFGATLVANEMNQDVLVFSALVFCAMVAVGVSSPAPDYAKAKYSAHVIMIEKSP 1019

```

Db 958 RFAGYVAHAKLMSEFEDLVFSAIVFGAMAVGVSSFAADYAKAKISAHHIMITEKTP 1017
QY 1020 IDTSHPGLKPNLEGNVFNENEVYPTRPDIPIVQLSLSEVKKQGLALVSGSGCKS 1079
Db 1018 IDSTSTGLMPNLEGNVFNENEVYPTRPDIPIVQLSLSEVKKQGLALVSGSGCKS 1077
QY 1080 TVVOLLERFVPLAGSVLIDGKEIKHLANTQWMLRAHGLGYSOBPIIFDCSIAENTAYGNS 1139
Db 1078 TVVOLLERFVPLAGSVLIDGKEIKHLANTQWMLRAHGLGYSOBPIIFDCSIAENTAYGNS 1137
QY 1140 RVVSHHEINQAKEANIHFIETLPKTYTRVGDGTOLSGGQKORIAIARALVROPHIL 1199
Db 1138 RVVSEIIVRAAKEANIHAFIESLPKTYSTKVGDKGTOLSGGQKORIAIARALVROPHIL 1197
QY 1200 LIDEATSAIDTESEKVVQALDKARGRPCYIARHRLSTIQNADLIYFONGKYEKGHTH 1259
Db 1198 LIDEATSAIDTESEKVVQALDKARGRPCYIARHRLSTIQNADLIYFONGKYEKGHTH 1257
QY 1260 QQLLAKGIFYSMVSYOAGAKR 1281
Db 1258 QQLLAKGIFYSMVSYOAGAKR 1279

RESULT 12
AAE18967
ID AAE18967 standard; Protein; 1280 AA.
AC AAE18967;
DT 21-MAY-2002 (first entry)
DE Human P-gp mutant, MDR1-L56P (8-A4).
KM Human; P-glycoprotein; vaccine; tumour; chemotherapeutic agent; cancer;
KN haematopoietic cell; cytotoxicity; cytostatic; P-gp; mutant; mutenin.
OS Homo sapiens.
OS Synthetic.
FH Key Location/Qualifiers
FT Misc-difference 56 /note="Wild type Leu substituted with Pro"
PN WC200210205-A2.
PD 07-FEB-2002.
PE 01-AUG-2001; 2001WO-US24560.
PR 01-AUG-2000; 2000US-222313P.
PA (UNII ) UNIV ILLINOIS FOUND.
PI Ruth A, Roninson I;
DR WPI; 2002-206182/26.
XX New human P-glycoprotein mutants for treating cancer, comprises an
PT ability to confer increased resistance to chemotherapeutic drugs
PT relative to wild type P-glycoprotein or P-glycoprotein having a glycine
PT to valine substitution
XX Claim 42; Fig 3; 47pp; English.
CC The present invention relates to an isolated mutant human P-glycoprotein
CC molecule. The invention is used as vaccine. An antibody specific for the
CC mutant P-glycoprotein is useful for determining whether a human tumour
CC sample comprises tumour cells expressing mutant human P-glycoprotein. A
CC recombinant expression construct comprising nucleic acid encoding mutant
CC P-glycoprotein is useful for constructing an increased amount of a
CC chemotherapeutic agent to an individual with cancer, by transducing ex
CC vivo haematopoietic cells from the individual with the recombinant

```

```

CC expression construct, reintroducing the transduced haematopoietic cells
CC into the individual and administering an increased amount of a
CC chemotherapeutic agent without consequent haematopoietic cytotoxicity.
CC The mutant P-glycoprotein is useful for treating and diagnosing cancer
CC and for screening compounds for the capacity to disrupt binding of
CC cytotoxic drugs with the mutant P-glycoprotein or to disrupt drug efflux
CC from cells expressing the mutant P-glycoprotein. The present sequence is
CC human P-glycoprotein (P-gp) mutant, MDR1-L56P (8-A4)
CC (multidrug resistance).
XX
SQ Sequence 1280 AA;
Query Match 90.5%; Score 5859; DB 23; Length 1280;
Best Local Similarity 90.7%; Pred. No. 0;
Matches 1163; Conservative 54; Mismatches 61; Indels 4; Gaps 4;
QY 1 MDEGGRKSA-EKNPMKGGKSKNEKEKPPYSTRFAMFRYSNMDRLVMYGTMAAI 59
Db 1 MDEGDRNGAKKKNFETKLNKSEK-DKEKKPYVSFMSFRYSNMDKLYMVVGTFAAI 59
QY 60 IHGAALPLMMLVFGNMTDSFANAGISRKTPPVIIINESITNNTQHTINHEEMTTAYY 119
Db 60 IHGAGLPMLMLVFGEMTDIFANAG-NLEDMSNTNSDINDTGFYKN-LEEDMTYAYY 117
QY 120 YSGIGAGVLAAYIOVSWCIIAAGROIIRKOFPHAIMROIGMFDVHDVDELNRLTD 179
Db 118 YSGIGAGVLAAYIOVSWCIIAAGROIIRKOFPHAIMROIGMFDVHDVDELNRLTD 177
QY 180 DYSKINEIGDKIGMFEFSIATFFTGFTGTRGKTLIVIAISPLVLSAIAIMAKIIS 239
Db 178 DYSKINEIGDKIGMFEFSIATFFTGFTGTRGKTLIVIAISPLVLSAIAIMAKIIS 237
QY 240 SFTDKELIAYAKAGAAVEVLAIRTVIAFGQKKELEERYKNLEAKGIGIKATIANI 299
Db 238 SFTDKELIAYAKAGAAVEVLAIRTVIAFGQKKELEERYKNLEAKGIGIKATIANI 297
QY 300 SIGAAFLITVASYALAFWYGSIVLSSEYSIGVLYFVSFVLGAISIGQASPIAFAN 359
Db 298 SIGAAFLITVASYALAFWYGSIVLSSEYSIGVLYFVSFVLGAISIGQASPIAFAN 357
QY 360 ARGAAVEIFKIIIDKPSIDSYSGKHKPDNIKGNLEKKNWFSEYPSRKVEKILKGINLV 419
Db 358 ARGAAVEIFKIIIDKPSIDSYSGKHKPDNIKGNLEKKNWFSEYPSRKVEKILKGINLV 417
QY 420 QSGOTVALVNGSGGKSTVQVMQRLYDPDGVACIDGODITINVRHLRETTGVVQEP 479
Db 418 QSGOTVALVNGSGGKSTVQVMQRLYDPDGVAVSDGODITINVRHLRETTGVVQEP 477
QY 480 VLFATTTAENIRYGRENVYMDIEKAVKEANAYDFIMKLPKFPDLYVGRGQSLSGGQK 539
Db 478 VLFATTTAENIRYGRENVYMDIEKAVKEANAYDFIMKLPKFPDLYVGRGQSLSGGQK 537
QY 540 RIAIARALVNRKILLIDEATSAIDTESAAYVOVALDKARKGRTTVIAHRLSTVRNDV 599
Db 538 RIAIARALVNRKILLIDEATSAIDTESAAYVOVALDKARKGRTTVIAHRLSTVRNDV 597
QY 600 IAGFDGVIVKGNHDELKKEKGYFKIYTMOTRNEIIEENATESSESSEDALEMSKD 659
Db 598 IAGFDGVIVKGNHDELKKEKGYFKIYTMOTRNEIIEENATESSESSEDALEMSKD 657
QY 660 SSGSLIKRSTRSRSHAPQGDRLGTKEDELENNPVPVSFMRILKLNSTEMPYPVYGFC 719
Db 658 SRSLSIKRSTRSRVSGQADRLSTKEADESLPPVSFMRILKLNSTEMPYPVYGFC 717
QY 720 AIINGLOPAFSLIRSLIIGITRDEDEPTKRONSMSVFLVGLIISITFFLQGTFF 779
Db 718 AIINGLOPAFSLIRSLIIGITRDEDEPTKRONSMSVFLVGLIISITFFLQGTFF 777
QY 780 GRAGETLKRLLRYVFRSLRQDVSWFDDPKNTGALTTRLANAAYKAGISRLAVIT 839
Db 778 GRAGETLKRLLRYVFRSLRQDVSWFDDPKNTGALTTRLANAAYKAGISRLAVIT 837
QY 840 QMIANLGIITISLITGQTLILLALVPIITAIAGVVEEMKMLSGQALDKRKELEGAKIA 899

```

```
Db 838 QNINNLGIIISFIYQMQLTLLLAIVPIIATAGVEKMLSGQALXKKKELEGAKIA 897
QY 900 TEALENFTVYSLIREQFEYMAQSLQVYRNSLRKAHLFVSFSTQAMTFYAGCF 959
Db 898 TEALENFTVYSLIQEQKEHMYAQSLQVYRNSLRKAHLFVFSTQAMTFYAGCF 957
QY 960 RFGALVANEPMNFQDVLVFSALVFGAMVGVSSFPADYAKAKSAAHVIMTIKSP 1019
Db 958 RFGALVANHKLMSFEDVLVFSAVVFGAMVGVSSFPADYAKAKSAAHVIMTIKSP 1017
QY 1020 IDSVPHLKPTLEGANTFENEVFYPTRPDIPLVGLSLLEVKGQTLALVSSCGSKS 1079
Db 1018 IDSISTEGMLPNTLEGANTFGEVFNPTRPDIPLVGLSLLEVKGQTLALVSSCGSKS 1077
QY 1080 TVVOLLERFYDPLAGSVLIDKEIKHLNVQMLRAHLGIVSEPLTFPCSTIAENIAYGDS 1139
Db 1078 TVVOLLERFYDPLAGSVLIDKEIKHLNVQMLRAHLGIVSEPLTFPCSTIAENIAYGDS 1137
QY 1140 RVVSHBEIQAQKAEKNIHHEITETLPEKYNTRVGDGTQLSGQKQRIATRAALVROPHIL 1199
Db 1138 RVVSOEIVRAKKAENIHAFETESLPNKSTKVGDKGTQLSGQKQRIATRAALVROPHIL 1197
QY 1200 LIDEATSALDTESEKRVQVQALDKAREGRTCIYIAHRLSTIONADLIYVFGNGKYEKGT 1259
Db 1198 LIDEATSALDTESEKRVQVQALDKAREGRTCIYIAHRLSTIONADLIYVFGNGKYEKGT 1257
QY 1260 QQLLAKGIFYFSMVSYQAGAKR 1281
Db 1258 QQLLAKGIFYFSMVSYQAGTKR 1279

RESULT 13
AAB81064
ID AAB81064 standard; Protein; 1280 AA.
XX
AC AAB81064:
XX
DT 25-JUN-2001 (first entry)
XX
DE Cynomologous monkey P-glycoprotein variant 1.
XX
KW Cynomologous monkey; P-glycoprotein; Pgp; multidrug transporter; MDR1;
XX
OS Macaca fascicularis.
XX
PN W0200123565-A1.
XX
PD 05-APR-2001.
XX
PE 28-SEP-2000; 2000MO-US26592.
XX
PR 28-SEP-1999; 99US-0156921.
XX
PR 12-OCT-1999; 99US-0158818.
XX
PA (GENT-) GENTEST CORP.
XX
PI Stocker PJ, Steimel-Crespi DT, Crespi CL;
XX
DR WPI: 2001-316136/33.
XX
DR N-PSDB: AAF86127.
XX
PT Novel isolated nucleic acid encoding cynomologous monkey P-glycoprotein
PT bivalability of compound and increasing pgp transporter activity in
PT cell
XX
PS Claim 9; Page 57-59; 84pp; English.
XX
CC This invention relates to a polynucleotide sequence encoding a
CC cynomologous monkey P-glycoprotein (Pgp), and an allelic variant of the
CC Pgp protein. Pgp, also known as multidrug transporter, MDR1 is a member
```

```
CC of the ABC transporter superfamily. The enzyme serves as an efflux pump
CC exporting small molecules across the cell membrane. The invention
CC includes a cynomologous monkey (Macaca fascicularis) Pgp coding sequence
CC and protein, and also that of an allelic variant. The Pgp polynucleotide
CC sequence is useful for increasing Pgp transporter activity in a cell.
CC Antisense sequences of the cDNA are useful for inhibiting Pgp transport
CC activity in a mammalian cell. They may also be used for increasing the
CC bioavailability of a drug. The present sequence represents the
CC cynomologous monkey P-glycoprotein.
XX
SQ Sequence 1280 AA:
Query Match 90.4%; Score 5857; DB 22; Length 1280;
Best Local Similarity 90.6%; Pred. No. 0;
Matches 1162; Conservative 59; Mismatches 57; Indels 4; Gaps 4;
QY 1 MDPEGGRKSAE-KNEFMKSKSKKNEKKEKPPVSTFAFRSNNLDRLYMVGTMAT 59
Db 1 MDPEGDRNGAEKKNFELNKSRR-KKERKPPVSFVSFRYSNNLDRLYMVGTLAAI 59
QY 60 IHGALPLIMLVFGMDTDFANAG-NICDGLALLTNSNTIDVYVWN-LEEDMTYAY 117
Db 60 IHGALPLIMLVFGMDTDFANAG-NICDGLALLTNSNTIDVYVWN-LEEDMTYAY 117
QY 120 YSIGAGVLYVAYIQVSPWCLAGRQILIKRQFFAHIMROEIGMFPVHVGELNRLTD 179
Db 118 YSIGAGVLYVAYIQVSPWCLAGRQILIKRQFFAHIMROEIGMFPVHVGELNRLTD 177
QY 180 DYSKINEGIGDKIGMFEHSIATFEFTGFTVGTFRCKMLTLVLAISPLVLSAAIMAKILS 239
Db 178 DYSKINEGIGDKIGMFEHSIATFEFTGFTVGTFRCKMLTLVLAISPLVLSAAIMAKILS 237
QY 240 SFTDKELLYAKAGAVAEVLAIKRTVIAFGQKKELEKRYKKNLEKAGIGIKATANI 299
Db 238 SFTDKELLYAKAGAVAEVLAIKRTVIAFGQKKELEKRYKKNLEKAGIGIKATANI 297
QY 300 SIGAFLLIYASALAFWGTSLVSSSEYSGIGVTFVFSVLIGAFSIGQSPSIEAFAN 359
Db 298 SIGAFLLIYASALAFWGTSLVSSSEYSGIGVTFVFSVLIGAFSIGQSPSIEAFAN 357
QY 360 ARGAAYEIRIIDNKSIDSYSKSGHKPDNIKGNLEFNKVPHSYPSREKVIILGNLKV 419
Db 358 ARGAAYEIRIIDNKSIDSYSKSGHKPDNIKGNLEFNKVPHSYPSREKVIILGNLKV 417
QY 420 QSGOTVALVNSGCGKSTVOLMQRLYDPTDGMNIDODIRITIVRLREITGVSGOE 479
Db 418 QSGOTVALVNSGCGKSTVOLMQRLYDPTDGMNIDODIRITIVRLREITGVSGOE 477
QY 480 VLFATTTAENIRYGRNVPMDIEKAVKANAYDFIMKLPKQFDPLVGERGAQSLSGQK 539
Db 478 VLFATTTAENIRYGRNVPMDIEKAVKANAYDFIMKLPKQFDPLVGERGAQSLSGQK 537
QY 540 RIATARALVBNPKILLDEATSALDTESEAVVOVALDKARKRTTIVIAHRLSTYRNADY 599
Db 538 RIATARALVBNPKILLDEATSALDTESEAVVOVALDKARKRTTIVIAHRLSTYRNADY 597
QY 600 IAGDDGVYIEKGNHDELMKEGITYFKLVYMTQNGEIELENATGSKESDALBMSPD 659
Db 598 IAGDDGVYIEKGNHDELMKEGITYFKLVYMTQNGEIELENATGSKESDALBMSPD 657
QY 660 SSSSLIKRSTRSRIRHAPQODRKLGTEDLNENVPVSWRLKLNSTEWYFVVGIFC 719
Db 658 SSSSLIKRSTRSRIRHAPQODRKLGTEDLNENVPVSWRLKLNSTEWYFVVGIFC 717
QY 720 AILINGGLAPASITIFSRTIGFTPDEDEPERKRONSNWESVLPLVGLISFITEFLOGFT 779
Db 718 AILINGGLAPAFVIFSKITIGFTNDAETKRONSNLFSLLPLVLIVGIFTEFLOGFT 777
QY 780 GRAGEITLTKRLRYWFRSMLRODVSWDFDPKNTGALTTRLANDAAQVGAIGSLAYTT 839
Db 778 GRAGEITLTKRLRYWFRSMLRODVSWDFDPKNTGALTTRLANDAAQVGAIGSLAYTT 837
QY 840 QNINNLGIIISFIYQMQLTLLLAIVPIIATAGVEKMLSGQALXKKKELEGAKIA 899
```

```
|||||
Db 838 ONIANLGTGIISLIYGMOLTLTLLAIYIILAGVVEKMLSGALDKKKEISGAGLIA 897
QY 900 TEAIENFRFVSLTRQKFEYMAAOSLOVPYNSLRKAIHICVNSTQAMMTSYACF 959
Db 898 TEAIENFRFVSLTRQKFEYMAAOSLOVPYNSLRKAIHICVNSTQAMMTSYACF 957
QY 960 REGAVIVANEFNFODVLVFSALVFGAMAGVSSFPADYAKAKVSAHYIMIEKSP 1019
Db 958 REGAVIVANESLMSFEDVLLVFSAVVFGAMAGVSSFPADYAKAKVSAHYIMIEKSP 1017
QY 1020 IDSYPHGAKPTLEGNTFNFNVTNPTRPDIYVGLSLSEVRKGOTLALVSGSGGKS 1079
Db 1018 IDSYPHGAKPTLEGNTFNFNVTNPTRPDIYVGLSLSEVRKGOTLALVSGSGGKS 1077
QY 1080 TVVOLLERYDPLAGSVLIDGKEIKLWOMLRALHIGVSEPTLFPSCIAENIAYGDS 1139
Db 1078 TVVOLLERYDPLAGSVLIDGKEIKLWOMLRALHIGVSEPTLFPSCIAENIAYGDS 1137
QY 1140 RVVSHETMQAAKEANIHFIETLPKRYNTRVGDGTQLSGGQKQRTAIRALVROPIL 1199
Db 1138 RVVSOEETVRAKEANIHAFIESLPNKYSTRVGDKGTQLSGGQKQRTAIRALVROPIL 1197
QY 1200 ILDEKTSALDPESEKVOEALDKAREGTCIVIAHRLSTIONADLIYVFGNGVKEHGT 1259
Db 1198 ILDEKTSALDPESEKVOEALDKAREGTCIVIAHRLSTIONADLIYVFGNGVKEHGT 1257
QY 1260 QOLLAQKGIYFSMVSQAGAKR 1281
Db 1258 QOLLAQKGIYFSMVSQAGAKR 1279

RESULT 14
AAE18968
ID AAE18968 standard; Protein: 1280 AA.
AC AAE18968;
XX
XX 21-MAY-2002 (first entry)
XX
XX Human P-gp mutant, MDR1-1190F/M197F (5-C3).
XX
XX Human; P-glycoprotein; vaccine; tumour; chemotherapeutic agent; cancer;
XX haematopoietic cell; cytotoxicity; cytosolic; P-gp; mutant; mutleln.
XX OS Homo sapiens.
XX OS Synthetic.
XX
XX Key Location/Qualifiers
XX FT MISC-difference 190 /note= "Wild type Ile substituted with Phe"
XX FT MISC-difference 197 /note= "Wild type Met substituted with Phe"
XX PN WO200210205-A2.
XX
XX 07-FEB-2002.
XX
XX 01-AUG-2001: 2001WO-US24560.
XX
XX 01-AUG-2000: 2000US-222313P.
XX
XX (UNIT ) UNIT ILLINOIS FOUND.
XX
XX Ruth A, Roninson I;
XX
XX MPI; 2002-206182/26.
XX
XX New human P-glycoprotein mutants for treating cancer. comprises an
XX ability to confer increased resistance to chemotherapeutic drugs
XX relative to wild type P-glycoprotein or P-glycoprotein having a glycine
XX to valine substitution
```

```
PS Claim 42: Fig 3; 47bp; English.
XX
CC The present invention relates to an isolated mutant human P-glycoprotein
CC molecule. The invention is used as vaccine. An antibody specific for the
CC mutant P-glycoprotein is useful for determining whether a human tumour
CC sample comprises tumour cells expressing mutant human P-glycoprotein. A
CC recombinant expression construct comprising nucleic acid encoding mutant
CC P-glycoprotein is useful for administering an increased amount of a
CC chemotherapeutic agent to an individual with cancer, by transducing ex
CC vivo haematopoietic cells from the individual with the recombinant
CC expression construct, reintroducing the transduced haematopoietic cells
CC into the individual and administering an increased amount of a
CC chemotherapeutic agent without consequent haematopoietic cytotoxicity.
CC The mutant P-glycoprotein is useful for treating and diagnosing cancer
CC and for screening compounds for the capacity to disrupt binding of
CC cytotoxic drugs with the mutant P-glycoprotein or to disrupt drug efflux
CC from cells expressing the mutant P-glycoprotein. The present sequence is
CC human P-glycoprotein (P-gp) mutant, MDR1-1190F/M197F (5-C3)
CC
XX
SQ Sequence 1280 AA;
Query Match 90.4%; Score 5856; DB 23; Length 1280;
Best Local Similarity 90.6%; Pred. No. 0;
Matches 1162; Conservative 54; Mismatches 62; Indels 4; Gaps 4;
QY 1 MDPEGGRKGS-A-EKNFWKMGKSKKNEKKRPVSTFAMFRYSNMIDRLYMLGTMAAI 59
Db 1 MDLEDNRNGAKKRNFFLNKRSK-DKEKKPVSYSFMSRYSNMIDRLYMLGTMAAI 59
QY 60 IHGAALPLMLVFNGMNTDPSFANAGISRNKTPVYINSTINNTQHLNHEEMTYAY 119
Db 60 IHGAGLPMLMVFEGEMTDIFRANAG-NLEDIMSNTNSDINDTGFFPM-LEEDMTYAY 117
QY 120 YSGIGAGVLAAYIQVSWFCIAAGRQILKTRKOFFHAIHQEIGWPDVHDVGLNTRLTD 179
Db 118 YSGIGAGVLAAYIQVSWFCIAAGRQILKTRKOFFHAIHQEIGWPDVHDVGLNTRLTD 177
QY 180 DVSKINGIGDKIGMFHSHATPTFTGTVGFTRGKILFVILIAISPYLGLSAIYNAKILS 239
Db 178 DVSKINGIGDKIGMFHSHATPTFTGTVGFTRGKILFVILIAISPYLGLSAIYNAKILS 237
QY 240 SFTDKELLAYAKAGAAVEEYLAIRIVYAGGOKKLEBYNNLEAKKIGIKKAITANI 299
Db 238 SFTDKELLAYAKAGAAVEEYLAIRIVYAGGOKKLEBYNNLEAKKIGIKKAITANI 297
QY 300 SIGAFLILITASYALAFWGTSLVLSSEYSIGQVLTVPFSVLGAFSIGQASPIEAFAN 359
Db 298 SIGAFLILITASYALAFWGTSLVLSSEYSIGQVLTVPFSVLGAFSIGQASPIEAFAN 357
QY 360 ARGAAVEIFKIIDNKPSIDSYSGSKHPNINGNLEFKNVHFSYSRKREYKILKLNKV 419
Db 358 ARGAAVEIFKIIDNKPSIDSYSGSKHPNINGNLEFRNVHFSYSRKREYKILKLNKV 417
QY 420 QSGQYVALVNGSGCKSTTVQMLQRLYDPTDGMVCIDGODITINVRHLREITGVVSE 479
Db 418 QSGQYVALVNGSGCKSTTVQMLQRLYDPTDGMVSGDQDITINVRHLREITGVVSE 477
QY 480 VLFATTLAENIRGRNVMDELEKAVKANAYDITMKLPKFDTLVGERGAQLSGGOK 539
Db 478 VLFATTLAENIRGRNVMDELEKAVKANAYDITMKLPKFDTLVGERGAQLSGGOK 537
QY 540 RIAIARALVRNPKIILDLDEATSLDTESEAVVQVALDKARKGRITIVIAHLSVRNADV 599
Db 538 RIAIARALVRNPKIILDLDEATSLDTESEAVVQVALDKARKGRITIVIAHLSVRNADV 597
QY 600 IAGFDGVIVEGNHDELMKEKGIYFKLVNQTGRNLELEENATGESKSESDALEMSPKD 659
Db 598 IAGFDGVIVEGNHDELMKEKGIYFKLVNQTGRNLEENADESKSEIDALEMSND 657
QY 660 SSSSLIKRSTRSIIHAPGODRKLTGKEDLNENVPVPSFRILKLNSTEMPVYVIFC 719
Db 658 SRSLSLIRKSTRSVRGSQADRLSTKEALDESIPVSEFWIRIMKLNLTMPVYVGVFC 717
```

```

QY 720 AIIINGIOPAPSIIFRSIIIGITFRDEDEPETKRONSNMFSVLVLGIIISFTFFELQGFTE 779
    |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 718 AIIINGIOPAPSIIFRSIIIGITFRDEDEPETKRONSNMFSVLVLGIIISFTFFELQGFTE 777
QY 780 GRAGEIITKRLRYMYFRSMLRQDVSMEDDKPKNTGALTTRLANDAAQYKAGISRLAVIT 839
    |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 778 GRAGEIITKRLRYMYFRSMLRQDVSMEDDKPKNTGALTTRLANDAAQYKAGISRLAVIT 837
QY 840 QINANGITIIISLIYGMQITLIIIAIPITAIAGVEMKMLSGOALKDKKELEGACKIA 899
    |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 838 QINANGITIIISLIYGMQITLIIIAIPITAIAGVEMKMLSGOALKDKKELEGACKIA 897
QY 900 TEAIEFRFVVSILTRBOKEFEYMYAASLOVPRNSLRKAHIFGVFSFTQAMMTSYAGCF 959
    |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 898 TEAIEFRFVVSILTRBOKEFEYMYAASLOVPRNSLRKAHIFGVFSFTQAMMTSYAGCF 957
QY 960 REGAYIVANEFNFOVLLVFSNAIVEGAMAVGVSSPAPDYAKAKVSAAHYIMITEKSPL 1019
    |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 958 REGAYIVANEFNFOVLLVFSNAIVEGAMAVGVSSPAPDYAKAKVSAAHYIMITEKSPL 1017
QY 1020 IDSYPHGLKPNTEGNVTENFENEVNYPTRPDIPLYOGISLEVKKGQTLALVSSGCGKS 1079
    |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 1018 IDSYPHGLKPNTEGNVTENFENEVNYPTRPDIPLYOGISLEVKKGQTLALVSSGCGKS 1077
QY 1080 TVVOLLERFYDPLAGSVLIDGKEIKHLNQWLRRAHLGIVSOEPIIFDCSIAMENIAYGDS 1139
    |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 1078 TVVOLLERFYDPLAGSVLIDGKEIKHLNQWLRRAHLGIVSOEPIIFDCSIAMENIAYGDS 1137
QY 1140 RYVSHHEIMQAKENANIHFIETLPKRYNTRVGDGKTQLSGQOKRIARALVRPQPHIL 1199
    |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 1138 RYVSHHEIMQAKENANIHFIETLPKRYNTRVGDGKTQLSGQOKRIARALVRPQPHIL 1197
QY 1200 LLDDEATSDLTSEKVVQDALDKAREGRICIVIAHFLSTIQNADLIYVQNKVHEHGH 1259
    |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 1198 LLDDEATSDLTSEKVVQDALDKAREGRICIVIAHFLSTIQNADLIYVQNKVHEHGH 1257
QY 1260 QQLLAKGITYFSMYSVOAGAKR 1281
    |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 1258 QQLLAKGITYFSMYSVOAGAKR 1279

RESULT 15
AAB81065
ID AAB81065 standard; protein; 1283 AA.
XX
AC AAB81065;
XX
DT 25-JUN-2001 (first entry)
XX
DE Cynomologous monkey P-glycoprotein variant 2.
XX
KW Cynomologous monkey; P-glycoprotein; PGP; multidrug transporter; MDRL;
    efflux pump.
XX
OS Macaca fascicularis.
XX
FH key location/Qualifiers
FT misc-difference 93..95 "An additional 3 amino acids are present compared
    to PGP variant AAB81064"
FT
FT
PN W0200123565-A1.
PD
PD 05-APR-2001.
PF 28-SEP-2000; 2000WO-US26592.
PR 28-SEP-1999; 99US-0156921.
PR 12-OCT-1999; 99US-0158818.
XX
XX (GENT-) GENTEST CORP.
XX
XX

```

```

PI Stocker PJ, Steimel-Crespi DT, Crespi CL;
XX WPI: 2001-316136/33.
DR N-PSDB; AAF86128.
XX
PT Novel isolated nucleic acid encoding cynomologous monkey P-glycoprotein
    (PGP) and homologous PGP polypeptides are useful for predicting
    bioavailability of compound and increasing PGP transporter activity in
    cell.
XX
PS Claim 9, page 65-68; 84pp; English.
XX
CC This invention relates to a polynucleotide sequence encoding a
    cynomologous monkey P-glycoprotein (PGP), and an allelic variant of the
    PGP protein. PGP, also known as multidrug transporter, MDRL is a member
    of the ABC transporter superfamily. The enzyme serves as an efflux pump
    exporting small molecules across the cell membrane. The invention
    includes a cynomologous monkey (macaca fascicularis) PGP coding sequence
    and protein, and also that of an allelic variant. The PGP polynucleotide
    sequence is useful for increasing PGP transporter activity in a cell.
    Antisense sequences of the cDNA are useful for inhibiting PGP transport
    activity in a mammalian cell. They may also be used for increasing the
    bioavailability of a drug. The present sequence represents the
    cynomologous monkey P-glycoprotein variant 2. The protein has an
    additional 3 amino acids when compared to PGP variant 1 (AAB81065).
XX
XX
SQ Sequence 1283 AA;
Query Match 90.4%; Score 5854.5; DB 22; Length 1283;
Best Local Similarity 90.3%; Pred. NO. 0;
Matches 1164; Conservative 57; Mismatches 53; Indels 15; Gaps 5;
QY 1 MDEEGGRKGSAAE-KNFMKMKSKSKNEKKRPVYSTAMFYSWMIDRLMLNGTMAI 59
    |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 1 MDEEGGRKGSAAEKKRPVYSTAMFYSWMIDRLMLNGTMAI 59
QY 60 IHGAALPIMLLVGNMTPSPANG-----ISRKTFPVYINESITNTQHFNIHLEEE 112
    |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 60 IHGAALPIMLLVGNMTPSPANG-----ISRKTFPVYINESITNTQHFNIHLEEE 113
QY 113 MTTVAAYYSIGAGVLAAYIQVSWFCLAAAGROIKIRKOPFAIMROEIGFVDHVDGE 172
    |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 114 MTTVAAYYSIGAGVLAAYIQVSWFCLAAAGROIKIRKOPFAIMROEIGFVDHVDGE 173
QY 173 LNTPLTDVSKINEIGDRTGMFEHSIATFTEGTYFTRGKRLTYLAIISPVLGLSAA 232
    |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 174 LNTPLTDVSKINEIGDRTGMFEHSIATFTEGTYFTRGKRLTYLAIISPVLGLSAA 233
QY 233 IWAIIISFTDKELLATAKAGAAVEYLAAIRTVIAFGQKKELERYKNLEBAKGIGIK 292
    |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 234 IWAIIISFTDKELLATAKAGAAVEYLAAIRTVIAFGQKKELERYKNLEBAKGIGIK 293
QY 293 KAITANISIGAAFLTYIASTALAFWGTSLVLSSEVSIGQVLFPSVLIAGAFSIGOASP 352
    |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 294 KAITANISIGAAFLTYIASTALAFWGTSLVLSSEVSIGQVLFPSVLIAGAFSIGOASP 353
QY 353 STEAPANRGAAYEFKTIIDNKPSIDSYSKSGKHPNINLEFFKNVHSPSRKEVKIL 412
    |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 354 STEAPANRGAAYEFKTIIDNKPSIDSYSKSGKHPNINLEFFKNVHSPSRKEVKIL 413
QY 413 KGLNLKQSGQTVALLVNGSGGKSTVOLMQLYDPTDGNVICDGDITRTINVRHLRETT 472
    |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 414 KGLNLKQSGQTVALLVNGSGGKSTVOLMQLYDPTDGNVICDGDITRTINVRHLRETT 473
QY 473 GVSQSEPVLFATTAENIRYGRVNTWDELKAVKANAADYIMKLPNKFDTLVGERGAO 532
    |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 474 GVSQSEPVLFATTAENIRYGRVNTWDELKAVKANAADYIMKLPNKFDTLVGERGAO 533
QY 533 LSGGOKRIATARALVRNPKTLLDEATSAIDTSEAVVOVALDKARKGRTTIVIAHRUS 592
    |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 534 LSGGOKRIATARALVRNPKTLLDEATSAIDTSEAVVOVALDKARKGRTTIVIAHRUS 593
QY 593 TVRNADVLAGPDGIVYKGNHDELMKEKGYIFKLVMTQTRNGEILEENATGESKESDA 652
    |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:

```



```

Db 594 TIRNADVLGPDGIVYKGNHDELMKEGITYKLYMOTAGNELLENADESKSETDT 653
QY 653 LEMSPKDSGSSLIKRSTRSRISHAFOGODRLGTREDLNENPPVPSFWRLIKLNSTENPY 712
Db 654 LEMSHDSGSSLIKRSTRSRVSGQDRLSTKEALDESIPVPSFWIRIMKLNTEMPY 713
QY 713 FVVGICAIINGLOPAFISITFSRIIGFTREDEPETKRONSMFSLFVLVGIISFTFE 772
Db 714 FVVGVCALINGLOPAFAVIFSKITIGFTREDDAETKRONSMFSLFLVLGIYSFTTF 773
QY 773 FLQGFTEGKAGELIKRLKRYMFRSMRLRODVSFDPDKNTGALTTRLANDAAYKGAIG 832
Db 774 FLQGFTEGKAGELIKRLKRYMFRSMRLRODVSFDPDKNTGALTTRLANDAAYKGAIG 833
QY 833 SRLAVITOMIANLGTGIIISLIYGWOLTLILLAIIVPIAIAGVEMKMLSGOALKDKKEL 892
Db 834 SRLAITONIANLGTGIIISLIYGWOLTLILLAIIVPIAIAGVEMKMLSGOALKDKKEL 893
QY 893 EGAGKIAATEAIEINFRTVVSILPREOKFEEMTAOSLOVPRNSLRKAHIFGVSPSTQAMMY 952
Db 894 EGAGKIAATEAIEINFRTVVSILPREOKFEEMTAOSLOVPRNSLRKAHIFGVSPSTQAMMY 953
QY 953 FSTAGCFRGAYIVANEFMNEODVILVFSALVFGAMAVGSSPADYAKAKVSAHAYIM 1012
Db 954 FSTAGCFRGAYIVANEFMNEODVILVFSALVFGAMAVGSSPADYAKAKVSAHAYIM 1013
QY 1013 IIEKSPILDSYSPHGLKPTLEGNTFENEVFNPTPRDIPVLOGLSLEVKKGOTLALVG 1072
Db 1014 IIEKTPILDSYSTEGIKPPTLEGNTFENEVFNPTPRDIPVLOGLSLEVKKGOTLALVG 1073
QY 1073 SSGGKSTVWOLLERFYDPLAGSVLDSQKEIKHLNOMLRANLGIYSOEPIIFDCSTIEN 1132
Db 1074 SSGGKSTVWOLLERFYDPLAGSVLDSQKEIKHLNOMLRANLGIYSOEPIIFDCSTIEN 1133
QY 1133 IAYGDNRRVVSHEIMQAARKEANIHHEITLPEKYNTRVGDKGTOLSGGQORAIARAL 1192
Db 1134 IAYGDNRRVVSHEIMQAARKEANIHHEITLPEKYNTRVGDKGTOLSGGQORAIARAL 1193
QY 1193 VROPHILLDEATSALDTSESEKVVQEAIDKARREGRTCIYAHRLSTIONADLIIVFONGK 1252
Db 1194 VROPHILLDEATSALDTSESEKVVQEAIDKARREGRTCIYAHRLSTIONADLIIVFONGK 1253
QY 1253 VKEHGTHOOLIAOKGIYFSMVSVOGAKR 1281
Db 1254 VKEHGTHOOLIAOKGIYFSMVSVOGAKR 1282

```

Search completed: December 9, 2002, 16:36:37
 Job time : 933 secs

GenCore version 5.1.3
Copyright (c) 1993 - 2002 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: December 9, 2002, 16:20:54 : Search time 142 Seconds

(without alignments)
146.524 Million cell updates/sec

Title: US-09-672-725C-2

Perfect score: 6477
Sequence: 1 MDPEGRKSGSAEKNFWMKMKGK.....LLAOKGTFMSVQAGAKR 1281

Scoring table:

BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 103943 seqs, 16242309 residues

Total number of hits satisfying chosen parameters: 103943

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 08
Maximum Match 100%

Listing first 45 summaries

Database : Published Applications_AA:*

- 1: /cgn2_6/ptodata/2/pubpaa/US08_NEW_PUB.pep.*
- 2: /cgn2_6/ptodata/2/pubpaa/PCIT_NEW_PUB.pep.*
- 3: /cgn2_6/ptodata/2/pubpaa/US06_NEW_PUB.pep.*
- 4: /cgn2_6/ptodata/2/pubpaa/US06_PUBCOMB.pep.*
- 5: /cgn2_6/ptodata/2/pubpaa/US07_NEW_PUB.pep.*
- 6: /cgn2_6/ptodata/2/pubpaa/US07_PUBCOMB.pep.*
- 7: /cgn2_6/ptodata/2/pubpaa/PCITUS_PUBCOMB.pep.*
- 8: /cgn2_6/ptodata/2/pubpaa/US08_PUBCOMB.pep.*
- 9: /cgn2_6/ptodata/2/pubpaa/US09_NEW_PUB.pep.*
- 10: /cgn2_6/ptodata/2/pubpaa/US09_PUBCOMB.pep.*
- 11: /cgn2_6/ptodata/2/pubpaa/US10_NEW_PUB.pep.*
- 12: /cgn2_6/ptodata/2/pubpaa/US10_PUBCOMB.pep.*
- 13: /cgn2_6/ptodata/2/pubpaa/US60_NEW_PUB.pep.*
- 14: /cgn2_6/ptodata/2/pubpaa/US60_PUBCOMB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	6420.5	99.1	1280	9	US-10-044-671-2
2	5861	90.5	1280	9	US-10-072-621-7
3	5861	90.5	1280	10	US-09-866-866A-2
4	5852	90.4	1280	10	US-09-866-866A-4
5	5636	87.0	1276	10	US-09-866-866A-8
6	5616	86.7	1272	10	US-09-769-097-4
7	5606	86.6	1276	10	US-09-866-866A-6
8	5296	81.8	1276	10	US-09-866-866A-6
9	3549	54.8	1195	10	US-09-873-409-5
10	3525.5	54.4	1222	10	US-09-873-409-5
11	3112.5	48.1	1058	10	US-09-873-409-4
12	2805.5	43.3	1275	10	US-09-873-409-4
13	2481	38.3	1334	9	US-09-749-340-6
14	2442	37.7	812	10	US-09-873-409-2
15	1950	30.1	1394	9	US-10-101-388-3
16	1870	28.9	659	10	US-09-873-409-1
17	1576	24.3	514	10	US-09-873-409-8
18	1552.5	24.0	541	10	US-09-873-409-7
19	836	12.9	582	10	US-09-815-242-14006

20	825	12.7	582	10	US-09-815-242-10115	Sequence 10115, A
21	820	12.7	766	9	US-10-072-621-6	Sequence 6, Appl1
22	816	12.6	766	10	US-09-795-693-17	Sequence 17, Appl1
23	815	12.6	587	10	US-09-815-242-10960	Sequence 10960, A
24	731	11.3	1261	10	US-09-759-143-538	Sequence 538, App
25	731	11.3	1261	10	US-09-780-669-538	Sequence 538, App
26	709.5	11.0	603	10	US-09-822-827-537	Sequence 537, App
27	709.5	11.0	603	10	US-09-815-242-5202	Sequence 5202, App
28	695.5	10.7	578	10	US-09-815-242-12501	Sequence 12501, A
29	687	10.6	748	9	US-09-870-759-41	Sequence 41, Appl1
30	672	10.4	640	9	US-09-970-059-9	Sequence 9, Appl1
31	664	10.3	571	10	US-09-815-242-10619	Sequence 10619, A
32	661	10.2	574	10	US-09-953-688A-1	Sequence 1, Appl1
33	661	10.2	836	10	US-09-953-688A-7	Sequence 7, Appl1
34	658	10.2	621	10	US-09-953-259-2	Sequence 2, Appl1
35	655	10.1	1228	10	US-09-759-143-537	Sequence 537, App
36	655	10.1	1228	10	US-09-780-669-537	Sequence 537, App
37	655	10.1	1228	10	US-09-822-827-537	Sequence 537, App
38	641.5	9.9	583	10	US-09-815-242-13458	Sequence 13458, A
39	636.5	9.8	339	9	US-09-764-884-25	Sequence 25, Appl1
40	618	9.5	353	9	US-09-764-884-33	Sequence 33, Appl1
41	596	9.2	1480	9	US-09-568-756-2	Sequence 2, Appl1
42	544	8.4	551	10	US-09-815-242-11394	Sequence 11394, A
43	528.5	8.2	400	10	US-09-765-272-190	Sequence 190, App
44	523.5	8.1	760	10	US-09-833-017-26	Sequence 26, Appl1
45	522.5	8.1	659	10	US-09-841-132-497	Sequence 497, App

ALIGNMENTS

```
RESULT 1
US-10-044-671-2
; Sequence 2, Application US/10044671
; Patent No. US2002017147A1
; GENERAL INFORMATION:
; APPLICANT: Washington State University Research Foundation
; APPLICANT: Mealey, Katrina
; TITLE OF INVENTION: Benthien, Steven
; FILE REFERENCE: MDRI VARIANTS AND METHODS FOR THEIR USE
; CURRENT APPLICATION NUMBER: US/10/044,671
; PRIOR FILING DATE: 2002-01-10
; PRIOR APPLICATION NUMBER: US 60/261,578
; PRIOR FILING DATE: 2001-01-12
; PRIOR APPLICATION NUMBER: US 60/314,829
; NUMBER OF SEQ ID NOS: 10
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 2
; LENGTH: 1280
; TYPE: PRT
; ORGANISM: Canis familiaris
US-10-044-671-2
Query Match          99.1% Score 6420.5; DB 9; Length 1280;
Best Local Similarity 99.3% Pred No. 0;
Matches 1272; Conservative 4; Mismatches 4; Indels 1; Gaps 1;
QY 1 MDPEGRKSGSAEKNFWMKMKGKSKKNEKKKPTVSFAMFRYSNNLDRILYMLVGTMAII 60
    |||
DB 1 MDPEGRKSGSAEKNFWMKMKGKSKK-EKREKKPTVSFAMFRYSNNLDRILYMLVGTMAII 59
    |||
QY 61 HGAALPLMLLVGNMDDPSANNGISRNKTFPVIINESTNTNTQHFNIHLEEMTVAAYY 120
    |||
DB 60 HGAALPLMLLVGNMDDPSANNGISRNKTFPVIINESTNTNTQHFNIHLEEMTVAAYY 119
    |||
QY 121 SGIGAGVLAAYIYVSWCLAGROILKIRKOFHAIKROEIGFVDVAGELWTRLTDD 180
    |||
DB 120 SGIGAGVLAAYIYVSWCLAGROILKIRKOFHAIKROEIGFVDVAGELWTRLTDD 179
    |||
QY 181 VKINIEGIGDKGMFFHSITATFFGFIYGFGRMKLTLLVLAISPVGLSNAIWAII 240
    |||
DB 181 VKINIEGIGDKGMFFHSITATFFGFIYGFGRMKLTLLVLAISPVGLSNAIWAII 240
    |||
```

```

Db 180 VSKINNEGIDKGVMPFOSIATFTTGTIVGFTPGMKLTVILAIISPVLGISAIAWAKILSS 239
Oy 241 FTDKELLAYAKAGAAVEEVLAAITVTIAFGGOKKLEERYKNLEEKAGIGIKKAITANIS 300
Db 240 FTDKELLAYAKAGAAVEEVLAAITVTIAFGGOKKLEERYKNLEEKAGIGIKKAITANIS 299
Oy 301 IGAATLLIYASVALAFWYGTSLVSSSEYSGOVLTFEFSVLIGAFSIGQSPSIEAFANA 360
Db 300 IGAATLLIYASVALAFWYGTSLVSSSEYSGOVLTFEFSVLIGAFSIGQSPSIEAFANA 359
Oy 361 RGAATFEIKIIDNKPSIDSYSGKHKPDNIKGNLEFNHVSPTSREKVKILKGLMLKYQ 420
Db 360 RGAATFEIKIIDNKPSIDSYSGKHKPDNIKGNLEFNHVSPTSREKVKILKGLMLKYQ 419
Oy 421 SGOYVALVNSCGCKSTTVQMLQRLYDPTDGMVCIDGODITINVRLHREITGVSGEPV 480
Db 420 SGOYVALVNSCGCKSTTVQMLQRLYDPTDGMVCIDGODITINVRLHREITGVSGEPV 479
Oy 481 LPATTAENIRYGRENVMTDEIEKAVKANAYDFIMKLPNKFDTLVGERGAKLSGGOKOR 540
Db 480 LPATTAENIRYGRENVMTDEIEKAVKANAYDFIMKLPNKFDTLVGERGAKLSGGOKOR 539
Oy 541 IATAALVNPVKIILIDEATSALDTESEAVVOVALDKARKKRTTIVIAHRLSTVRNADVI 600
Db 540 IATAALVNPVKIILIDEATSALDTESEAVVOVALDKARKKRTTIVIAHRLSTVRNADVI 599
Oy 601 AGPDGVIEKGNHDELMKEKGIYFKLYVMQTRGNEIELEENATGESKESDALEMSPKDS 660
Db 600 AGPDGVIEKGNHDELMKEKGIYFKLYVMQTRGNEIELEENATGESKESDALEMSPKDS 659
Oy 661 GSSLIKRSTRRSIAHPOGODRKLGTKEDLNENPVPSFWRLIKNSTEMPEYFVVGIFCA 720
Db 660 GSSLIKRSTRRSIAHPOGODRKLGTKEDLNENPVPSFWRLIKNSTEMPEYFVVGIFCA 719
Oy 721 IINGGLOPASFIFSRITIGITRDEDEPETKRONSNMFSLVLVLGIIISITIFFLOGFTGG 780
Db 720 IINGGLOPASFIFSRITIGITRDEDEPETKRONSNMFSLVLVLGIIISITIFFLOGFTGG 779
Oy 781 KAGEIILTKRLRYVFRSM.LRODVSFMDPKNTTGALTTRFLANDAQVGAIGSLRAYTQ 840
Db 780 KAGEIILTKRLRYVFRSM.LRODVSFMDPKNTTGALTTRFLANDAQVGAIGSLRAYTQ 839
Oy 841 NIANLGGIILSLITGMQLLILALVPIIATAGVEMKLSGQALDKDELEGAKIAT 900
Db 840 NIANLGGIILSLITGMQLLILALVPIIATAGVEMKLSGQALDKDELEGAKIAT 899
Oy 901 EAIENFTVVSITREOKFEFYMAQSLQVPYRNSLRKAHIFGVSFSTTQAMFTSYAGCER 960
Db 900 EAIENFTVVSITREOKFEFYMAQSLQVPYRNSLRKAHIFGVSFSTTQAMFTSYAGCER 959
Oy 961 FGATLVANEPKNDVLVLSAIVFGMAVGOVSSFAPDYAKAKVSAAHVIMIIKESPLI 1020
Db 960 FGATLVANEPKNDVLVLSAIVFGMAVGOVSSFAPDYAKAKVSAAHVIMIIKESPLI 1019
Oy 1021 DSYSPHGLKPTLTGANTFENEVFNPTRPDIPIVLOGLSLEVKKGOTLALVSSGGCKST 1080
Db 1020 DSYSPHGLKPTLTGANTFENEVFNPTRPDIPIVLOGLSLEVKKGOTLALVSSGGCKST 1079
Oy 1081 VVOLLEFRYDPLAGSVLIDGKEIKHLNWCMLRAHLGIVSGEPILFPSCIAENIAYGDNR 1140
Db 1080 VVOLLEFRYDPLAGSVLIDGKEIKHLNWCMLRAHLGIVSGEPILFPSCIAENIAYGDNR 1139
Oy 1141 VVSHEIQAQKANIHHFETLPEKYNTRYGDKGTOLSGGOKORIAIARALVROPHILL 1200
Db 1140 VVSHEIQAQKANIHHFETLPEKYNTRYGDKGTOLSGGOKORIAIARALVROPHILL 1199
Oy 1201 IDEATSALDTESEKVVQALDKARBGTCTVIAHRLSTIONADLIYVFOGKXKHEGTHQ 1260
Db 1200 IDEATSALDTESEKVVQALDKARBGTCTVIAHRLSTIONADLIYVFOGKXKHEGTHQ 1259
Oy 1261 OLLAOKGIYFSMVYQAGAKR 1281
Db 1260 OLLAOKGIYFSMISVQAGAKR 1280

```

```

RESULT 2
US-10-072-621-7
: Sequence 7, Application US/10072621
: Patient No. US20020169137A1
: GENERAL INFORMATION:
: APPLICANT: Reiner, Peter B.
: APPLICANT: Connop, Bruce P.
: APPLICANT: Pollard, Michelle
: TITLE OF INVENTION: REGULATION OF AMYLOID PRECURSOR PROTEIN EXPRESSION
: TITLE OF INVENTION: BY MODIFICATION OF ABC TRANSPORTER EXPRESSION OR ACTIVITY
: FILE REFERENCE: 100103.402
: CURRENT APPLICATION NUMBER: US/10/072.621
: CURRENT FILING DATE: 2002-02-08
: NUMBER OF SEQ ID NOS: 10
: SOFTWARE: FastSeq for Windows Version 4.0
: SEQ ID NO 7
: LENGTH: 1280
: TYPE: PRT
: ORGANISM: Homo sapiens
US-10-072-621-7

```

```

Query Match 90.5% Score 5861: DB 9; Length 1280:
Best Local Similarity 90.7% Pred. No. 0:
Matches 1163: Conservative 55; Mismatches 60; Indels 4; Gaps 4:

Oy 1 MDPEGGRGSA-EKNFMKMGKSKKNEKKEKPTVSTFAMFRYSNMJDRULYLVGTMAI 59
Db 1 MDPEGGRGSAKKKPFLLNKSEK-DKKEKKPIVSVSMFRYSNMJDKLYMVGTMAI 59
Oy 60 IHGALPLMMLVFGNMVDSFANAGISRNKTPVILINESITNNQHFNLHEEMTVAYY 119
Db 60 IHGAGPLMMLVFGEMTDFANAG-NLEDLMSNTNRSDINDTDFEM-LEEDMTRAYY 117
Oy 120 YSGTGGVVAAYIOYSEFCLAGROILKIRKOPFAHIMROEIGMFDVHDVGEINRLTD 179
Db 118 YSGTGGVVAAYIOYSEFCLAGROILKIRKOPFAHIMROEIGMFDVHDVGEINRLTD 177
Oy 180 DYSKINEGIDKIGMEFHSIATFTTGTIVGFTPGMKLTVILAIISPVLGISAIAWAKILS 239
Db 178 DYSKINEGIDKIGMEFHSIATFTTGTIVGFTPGMKLTVILAIISPVLGISAIAWAKILS 237
Oy 240 FTDKELLAYAKAGAAVEEVLAAITVTIAFGGOKKLEERYKNLEEKAGIGIKKAITANI 299
Db 238 FTDKELLAYAKAGAAVEEVLAAITVTIAFGGOKKLEERYKNLEEKAGIGIKKAITANI 297
Oy 300 SIGAATLLIYASVALAFWYGTSLVSSSEYSGOVLTFEFSVLIGAFSIGQSPSIEAFAN 360
Db 298 SIGAATLLIYASVALAFWYGTSLVSSSEYSGOVLTFEFSVLIGAFSIGQSPSIEAFAN 357
Oy 360 ARGAAVEIFKIIDNKPSIDSYSGKHKPDNIKGNLEFNHVSPTSREKVKILKGLMLKY 420
Db 358 ARGAAVEIFKIIDNKPSIDSYSGKHKPDNIKGNLEFNHVSPTSREKVKILKGLMLKY 417
Oy 420 OSGOTVALVNSCGCKSTTVQMLQRLYDPTDGMVCIDGODITINVRLHREITGVSGEPV 480
Db 418 OSGOTVALVNSCGCKSTTVQMLQRLYDPTDGMVCIDGODITINVRLHREITGVSGEPV 477
Oy 480 VLFATTAENIRYGRENVMTDEIEKAVKANAYDFIMKLPNKFDTLVGERGAKLSGGOKOR 540
Db 478 VLFATTAENIRYGRENVMTDEIEKAVKANAYDFIMKLPNKFDTLVGERGAKLSGGOKOR 537
Oy 540 RIAARALVNPVKIILIDEATSALDTESEAVVOVALDKARKKRTTIVIAHRLSTVRNADVI 600
Db 538 RIAARALVNPVKIILIDEATSALDTESEAVVOVALDKARKKRTTIVIAHRLSTVRNADVI 597
Oy 600 IAGPDGVIEKGNHDELMKEKGIYFKLYVMQTRGNEIELEENATGESKESDALEMSPKDS 660
Db 598 IAGPDGVIEKGNHDELMKEKGIYFKLYVMQTRGNEIELEENATGESKESDALEMSPKDS 657
Oy 660 GSSLIKRSTRRSIAHPOGODRKLGTKEDLNENPVPSFWRLIKNSTEMPEYFVVGIFC 720
Db 658 GSSLIKRSTRRSIAHPOGODRKLGTKEDLNENPVPSFWRLIKNSTEMPEYFVVGIFC 719

```

```
Db 658 SSSLIKRSTRRSVRSQADRLSTKEALDESIPPVSFWRLMKLNTLTEMPYVGVGC 717
Qy 720 AIIINGLOPAFISIFSRIGIFTRDEDEPETKRONSNFSLVFLVIGISITFFLOGFTF 779
Db 718 AIIINGLOPAFAIFSKITIGFTRIDDEPETKRONSNFSLFLALGIIISFTIFFLOGFTF 777
Qy 780 GKAGEILLKRLRYVFRSMRLRODVSWFDDPKNTGALTTRLANDAAOVGAIGSLRLAVIT 839
Db 778 GKAGEILLKRLRYVFRSMRLRODVSWFDDPKNTGALTTRLANDAAOVGAIGSLRLAVIT 837
Qy 840 QNIANLGTGIIISLYGMOQLTLLLAIVPIIATAGVEMKMSGQALKDKKELEGAGKTA 899
Db 838 QNIANLGTGIIISLYGMOQLTLLLAIVPIIATAGVEMKMSGQALKDKKELEGAGKTA 897
Qy 900 TEALENFRVVSILTRQEKFEYMAQSLQVYRNSLRKAHIFGVSTIQAMMYFYACGF 959
Db 898 TEALENFRVVSILTRQEKFEYMAQSLQVYRNSLRKAHIFGVSTIQAMMYFYACGF 957
Qy 960 RFGAYLVANFNFQDVLVFSALVFGAMAVQVSSFPADYAKAKVSAHVIMIIEKSP 1019
Db 958 RFGAYLVANFNFQDVLVFSALVFGAMAVQVSSFPADYAKAKVSAHVIMIIEKSP 1017
Qy 1020 IDSYSBGLKPNLTLEGNVTENEVEVNYPTRPDIPVLOGISLEVKKGQTLALVSSGCGKS 1079
Db 1018 IDSYSBGLKPNLTLEGNVTENEVEVNYPTRPDIPVLOGISLEVKKGQTLALVSSGCGKS 1077
Qy 1080 TVVOLLERYPDLASVLLDGEKEIKHLNVOMLRHLGIVSOEPIIFDCSIENIAYGDN 1139
Db 1078 TVVOLLERYPDLASVLLDGEKEIKHLNVOMLRHLGIVSOEPIIFDCSIENIAYGDN 1137
Qy 1140 RVVSHIEIMQAKEANIHFIETLEPKYNTRVGDKGTOLSGGOKRIATARALVROPHIL 1199
Db 1138 RVVSGEIVRAKKEANIHFIETLEPKYNTRVGDKGTOLSGGOKRIATARALVROPHIL 1197
Qy 1200 LIDEATSALDTESEKVOYDALKAREGRTCIIVIAHRLSTIONMADLIYFONGKXKEGHT 1259
Db 1198 LIDEATSALDTESEKVOYDALKAREGRTCIIVIAHRLSTIONMADLIYFONGKXKEGHT 1257
Qy 1260 QOLLAOKGIYFSMVSVQAGAR 1281
Db 1258 QOLLAOKGIYFSMVSVQAGARK 1279

RESULT 3
US-09-866-866A-2
; Sequence 2, Application US/09866866A
; Patent No. US20020102244A1
; GENERAL INFORMATION:
; APPLICANT: Sorrentino, Brian
; TITLE OF INVENTION: A Method of Identifying and/or Isolating Stem Cells
; FILE REFERENCE: 1340-1-021CIP2
; CURRENT APPLICATION NUMBER: US/09/866, 866A
; PRIOR APPLICATION NUMBER: 2001-08-30
; PRIOR FILING DATE: 2000-05-31
; PRIOR APPLICATION NUMBER: PCT/US99/11825
; PRIOR FILING DATE: 1999-05-27
; PRIOR APPLICATION NUMBER: 60/086, 988
; PRIOR FILING DATE: 1998-05-28
; NUMBER OF SEQ ID NOS: 27
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 2
; LENGTH: 1280
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-866-866A-2
```

Query Match 90.5%; Score 5861; DB 10; Length 1280;
Best Local Similarity 90.7%; Pred. No. 0;
Matches 1163; Conservative 55; Mismatches 60; Indels 4; Gaps 4;

Qy 1 MDEGGRKGA-EKNWKGKSKKNEKKEKPYSTAMFRTYSNMWLDRLIYLVGTMAAI 59

```
Db 1 MDLGDGRNGAKKKNFKLNKSEK -DKKEKPTVSVFSERKSNMWDLIYVGTMAAI 59
Qy 60 IHGAALPLMIVYENMMDSPANNAGISNKTPEYIINESTNNTOHFINHLEEMTYAY 119
Db 60 IHGAGLPLMIVYENMMDSPANNAGISNKTPEYIINESTNNTOHFINHLEEMTYAY 117
Qy 120 YSGIAGVLAAYLOVSWFLAAGROLTKIRKOFPAIRKROEIGMDVDVDELNTRLTD 179
Db 118 YSGIAGVLAAYLOVSWFLAAGROLTKIRKOFPAIRKROEIGMDVDVDELNTRLTD 177
Qy 180 DVSKINGIDCKIGMFPHSIATFTFTGIVFTGKMLTVLIIAISVGLSAIAWAKLIS 239
Db 178 DVSKINGIDCKIGMFPHSIATFTFTGIVFTGKMLTVLIIAISVGLSAIAWAKLIS 237
Qy 240 SFTDKELLAAYAKAGAAVEYLARTVIAFGQKLEKRYNKNLEBANGIGIKRATANI 299
Db 238 SFTDKELLAAYAKAGAAVEYLARTVIAFGQKLEKRYNKNLEBANGIGIKRATANI 297
Qy 300 SIGAPELLIYASALAFWYGTSLVLSSEYSIGQVLTFFSVLIGAFSIGOASPTAEFAN 359
Db 298 SIGAPELLIYASALAFWYGTSLVLSSEYSIGQVLTFFSVLIGAFSIGOASPTAEFAN 357
Qy 360 ARGAAYEIKIIDNKSIDSYSKSGHKPUNIKGNLEPKNVHSPYSRREKVALGLNLKV 419
Db 358 ARGAAYEIKIIDNKSIDSYSKSGHKPUNIKGNLEPKNVHSPYSRREKVALGLNLKV 417
Qy 420 QSGQVVALVNGSGCGKSTTVOLMORLYDPTDGMVINDQDRIITNVRRLREITGVSOEP 479
Db 418 QSGQVVALVNGSGCGKSTTVOLMORLYDPTDGMVINDQDRIITNVRRLREITGVSOEP 477
Qy 480 VLEATTIANITRYGENTYMDIEKAYEANAAYDFIMKLPKFDVLVGERCAQLSGGOK 539
Db 478 VLEATTIANITRYGENTYMDIEKAYEANAAYDFIMKLPKFDVLVGERCAQLSGGOK 537
Qy 540 RIAIARALVNRKILLDEATSALDTESEAVYVALDKARGRTTIVIAHRLSTVRNADY 599
Db 538 RIAIARALVNRKILLDEATSALDTESEAVYVALDKARGRTTIVIAHRLSTVRNADY 597
Qy 600 IAGFDGYIVKGNHDELMKEKGIYFKLVYMTQGNIELEENATGESKSDALEMSPKD 659
Db 598 IAGFDGYIVKGNHDELMKEKGIYFKLVYMTQGNIELEENATGESKSDALEMSPKD 657
Qy 660 SSSLIKRSTRRSVRSQADRLSTKEALDESIPPVSFWRLMKLNTLTEMPYVGVGC 719
Db 658 SSSLIKRSTRRSVRSQADRLSTKEALDESIPPVSFWRLMKLNTLTEMPYVGVGC 717
Qy 720 AIIINGLOPAFISIFSRIGIFTRDEDEPETKRONSNFSLVFLVIGISITFFLOGFTF 779
Db 718 AIIINGLOPAFAIFSKITIGFTRIDDEPETKRONSNFSLFLALGIIISFTIFFLOGFTF 777
Qy 780 GKAGEILLKRLRYVFRSMRLRODVSWFDDPKNTGALTTRLANDAAOVGAIGSLRLAVIT 839
Db 778 GKAGEILLKRLRYVFRSMRLRODVSWFDDPKNTGALTTRLANDAAOVGAIGSLRLAVIT 837
Qy 840 QNIANLGTGIIISLYGMOQLTLLLAIVPIIATAGVEMKMSGQALKDKKELEGAGKTA 899
Db 838 QNIANLGTGIIISLYGMOQLTLLLAIVPIIATAGVEMKMSGQALKDKKELEGAGKTA 897
Qy 900 TEALENFRVVSILTRQEKFEYMAQSLQVYRNSLRKAHIFGVSTIQAMMYFYACGF 959
Db 898 TEALENFRVVSILTRQEKFEYMAQSLQVYRNSLRKAHIFGVSTIQAMMYFYACGF 957
Qy 960 RFGAYLVANFNFQDVLVFSALVFGAMAVQVSSFPADYAKAKVSAHVIMIIEKSP 1019
Db 958 RFGAYLVANFNFQDVLVFSALVFGAMAVQVSSFPADYAKAKVSAHVIMIIEKSP 1017
Qy 1020 IDSYSBGLKPNLTLEGNVTENEVEVNYPTRPDIPVLOGISLEVKKGQTLALVSSGCGKS 1079
Db 1018 IDSYSBGLKPNLTLEGNVTENEVEVNYPTRPDIPVLOGISLEVKKGQTLALVSSGCGKS 1077
Qy 1080 TVVOLLERYPDLASVLLDGEKEIKHLNVOMLRHLGIVSOEPIIFDCSIENIAYGDN 1139
```

```
Db 1078 TVVOLLERFYDPLAGVLLDGEIKRLNOMLRAHGIYSOEPIFDCSIENIAVGDNS 1137
Oy 1140 RVVSHHEIMOAKRANIHHEITLPEKYNTRVGDKGTOLSGGOKORIAIARALVROPAIL 1199
    |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Db 1138 RVVSGEIVRAKAKRANHAFIESLPNKYSTKVGDKGTOLSGGOKORIAIARALVROPAIL 1197
Oy 1200 LLDENATSDTESEKVVQVQALDKAREGRTCIYIAHRLSTIONADLIYVQNGVKEHGT 1259
    |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Db 1198 LLDENATSDTESEKVVQVQALDKAREGRTCIYIAHRLSTIONADLIYVQNGVKEHGT 1257
Oy 1260 QOLLAOKGIYFSMVSVOAGAKR 1281
    |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Db 1258 QOLLAOKGIYFSMVSVOAGAKR 1279

RESULT 4
US-09-866-866A-4
; Sequence 4, Application US/09866866A
; Patent No. US20020102244A1
; GENERAL INFORMATION:
; APPLICANT: Sorrentino, Brian
; APPLICANT: Schuetz, John
; TITLE OF INVENTION: A Method of Identifying and/or Isolating Stem Cells
; FILE REFERENCE: 1340-1-021CIP2
; CURRENT APPLICATION NUMBER: US/09/866, 866A
; CURRENT FILING DATE: 2001-08-30
; PRIOR APPLICATION NUMBER: 09/584, 586
; PRIOR FILING DATE: 2000-05-31
; PRIOR APPLICATION NUMBER: PCT/US99/11825
; PRIOR FILING DATE: 1999-05-27
; PRIOR APPLICATION NUMBER: 60/086, 988
; PRIOR FILING DATE: 1998-05-28
; NUMBER OF SEQ ID NOS: 27
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 4
; LENGTH: 1280
; TYPE: PRF
; ORGANISM: Homo sapiens
US-09-866-866A-4

Query Match          90.4%; Score 5852; DB 10; Length 1280;
Best Local Similarity 90.6%; Pred. No. 0;
Matches 1162; Conservative 55; Mismatches 61; Indels 4; Gaps 4;

Oy 1 MDPEGGRKGSN-EKNFWKMGKSKKNEKKEKPTVSTFAMFRYSNMWLDLYLNGTMAI 59
    |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Db 1 MDEGDNRNGGAKKKNNKNNKSEK-DKKEKKPTVSFMSFRTSNWLDKLYVVGTLAAT 59
Oy 60 IHGAALPLMLVFGNMTDFRANGISNKTFFVITNESITNTNTOHFINHLEEMTYAYY 119
    |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Db 60 IHGAGLPMLVFGEMTDIFRANG-NLEDLMSNITNRSNDNIGTFPM-LIEDMTRFAYY 117
Oy 120 YSGIGAGVLAAYIQVSWFCLAGROILKIRKOFHAIMROETIGMFDVHVGELNTRLD 179
    |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Db 118 YSGIGAGVLAAYIQVSWFCLAGROILKIRKOFHAIMROETIGMFDVHVGELNTRLD 177
Oy 180 DVSKINIEGDKIGMFRHSIATFTFTGTVGTRKMKLTIVILASPLGISAIAIMAKIIS 239
    |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Db 178 DVSKINIEYDKIGMFRHSIATFTFTGTVGTRKMKLTIVILASPLGISAIAIMAKIIS 237
Oy 240 SPTDEKLAAYKAGAVAEVLAIRTYIAFGGOKKELERYNKNLEAKGIGIKKAITANT 299
    |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Db 238 SPTDEKLAAYKAGAVAEVLAIRTYIAFGGOKKELERYNKNLEAKGIGIKKAITANT 297
Oy 300 SIGAFLIYASYALAFMYGTSLVISEYSGIYLVYFVSYLGAISIGASPIEAFAN 359
    |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Db 298 SIGAFLIYASYALAFMYGTSLVISEYSGIYLVYFVSYLGAISIGASPIEAFAN 357
Oy 360 AGAAVEIFKTIIDNKPSIDSYSGSKGRPDNIKNIIEFNHVSYPSSREKVKILGLMLKY 419
    |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Db 358 AGAAVEIFKTIIDNKPSIDSYSGSKGRPDNIKNIIEFNHVSYPSSREKVKILGLMLKY 417
Oy 420 OSGQTVIALVGNCGCKSTTVOLMORLYDPDGMVCIDGQDIRTINVRKHLREITGVSOEP 479
```

```
Db 418 OSGQTVIALVGNCGCKSTTVOLMORLYDPDGMVSVDOGDIRTINVRKHLREITGVSOEP 477
    |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Oy 480 VLEFATTAENIRYGREVYVWDEIEKAVKEANAYDFIMKLPNFDLVGSGAQOLSGGOK 539
    |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Db 478 VLEFATTAENIRYGREVYVWDEIEKAVKEANAYDFIMKLPNFDLVGSGAQOLSGGOK 537
Oy 540 RIATARALVNPRIKLLIDEATSDTESEAVVOALDKARKGRTCIYIAHRLSTIONAD 599
    |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Db 538 RIATARALVNPRIKLLIDEATSDTESEAVVOALDKARKGRTCIYIAHRLSTIONAD 597
Oy 600 IAGPDGVIVKGNHDELMKEKGIYFKLYMTQTRGNEIELENATGSESDALEMSPD 659
    |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Db 598 IAGPDGVIVKGNHDELMKEKGIYFKLYMTQTRGNEIELENATGSESDALEMSPD 657
Oy 660 SSSSLIKRRSTRSRIHAPOGDRKLTGEDLINEVNPVPSWRILKLNSTPEYVVGIFC 719
    |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Db 658 SSSSLIKRRSTRSRIHAPOGDRKLTGEDLINEVNPVPSWRILKLNSTPEYVVGIFC 717
Oy 720 AIIINGLOPAPSIIFSRIGIFTRDEDEPFRKONSNEFVLVLGIISFTFELQGFTE 779
    |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Db 718 AIIINGLOPAPSIIFSRIGIFTRDEDEPFRKONSNEFVLVLGIISFTFELQGFTE 777
Oy 780 GKAGEILTRKLYVFRSMLRODVSWEDEPKNTGALTTRLANDAAYKALISRLAVIT 839
    |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Db 778 GKAGEILTRKLYVFRSMLRODVSWEDEPKNTGALTTRLANDAAYKALISRLAVIT 837
Oy 840 ONIANIGTITISLYGQOLITLILAIYPIAIAIGVEMKMLSGQALDKKELEGACKIA 899
    |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Db 838 ONIANIGTITISLYGQOLITLILAIYPIAIAIGVEMKMLSGQALDKKELEGACKIA 897
Oy 900 TEAIENTRVSLTRBOKREYVYVYVYVYVYVYVYVYVYVYVYVYVYVYVYVYVYVYV 959
    |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Db 898 TEAIENTRVSLTRBOKREYVYVYVYVYVYVYVYVYVYVYVYVYVYVYVYVYVYVY 957
Oy 960 REGAVLVANEFKMFQDVLVFSAYIFGMAVGOVSSFPADYAKKISAHHIMIEKTP 1019
    |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Db 958 REGAVLVANEFKMFQDVLVFSAYIFGMAVGOVSSFPADYAKKISAHHIMIEKTP 1017
Oy 1020 IDSYPHGLKPNTEGNVTENEVENYPTRPDIYVLOGISLEVKKGOTLALVSSGCKS 1079
    |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Db 1018 IDSYPHGLKPNTEGNVTENEVENYPTRPDIYVLOGISLEVKKGOTLALVSSGCKS 1077
Oy 1080 TVVOLLERFYDPLAGVLLDGEIKRLNOMLRAHGIYSOEPIFDCSIENIAVGDNS 1139
    |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Db 1078 TVVOLLERFYDPLAGVLLDGEIKRLNOMLRAHGIYSOEPIFDCSIENIAVGDNS 1137
Oy 1140 RVVSHHEIMOAKRANIHHEITLPEKYNTRVGDKGTOLSGGOKORIAIARALVROPAIL 1199
    |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Db 1138 RVVSGEIVRAKAKRANHAFIESLPNKYSTKVGDKGTOLSGGOKORIAIARALVROPAIL 1197
Oy 1200 LLDENATSDTESEKVVQVQALDKAREGRTCIYIAHRLSTIONADLIYVQNGVKEHGT 1259
    |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Db 1198 LLDENATSDTESEKVVQVQALDKAREGRTCIYIAHRLSTIONADLIYVQNGVKEHGT 1257
Oy 1260 QOLLAOKGIYFSMVSVOAGAKR 1281
    |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Db 1258 QOLLAOKGIYFSMVSVOAGAKR 1279

RESULT 5
US-09-866-866A-8
; Sequence 8, Application US/09866866A
; Patent No. US20020102244A1
; GENERAL INFORMATION:
; APPLICANT: Sorrentino, Brian
; APPLICANT: Schuetz, John
; TITLE OF INVENTION: A Method of Identifying and/or Isolating Stem Cells
; FILE REFERENCE: 1340-1-021CIP2
; CURRENT APPLICATION NUMBER: US/09/866, 866A
; CURRENT FILING DATE: 2001-08-30
; PRIOR APPLICATION NUMBER: 09/584, 586
; PRIOR FILING DATE: 2000-05-31
```


; PRIOR APPLICATION NUMBER: PCT/US99/11825
 ; PRIOR FILING DATE: 1999-05-27
 ; PRIOR APPLICATION NUMBER: 60/086, 988
 ; PRIOR FILING DATE: 1998-05-28
 ; NUMBER OF SEQ ID NOS: 27
 ; SOFTWARE: PatentIn version 3.0
 ; SEQ ID NO 8
 ; LENGTH: 1276
 ; TYPE: PRT
 ; ORGANISM: Mus musculus
 us-09-866-866A-8

Query Match 87.0%; Score 5636; DB 10; Length 1276;
 Best Local Similarity 87.0%; Pred. No. 0;
 Matches 115; Conservative 75; Mismatches 84; Indels 8; Gaps 4;

QY 1 MDEGGRKSAEKNFMKSKKKKKKKPTVSFAMFRKSNMLDRLYMLVGTMAII 60
 DB 1 MELEEDLKGADKNFSKMGKSKK-EKKKKPAVSVLTFMFRAGWLDRLYMLVGTMAII 59
 QY 61 HGAALPLMLVFGNMTDSFANNG-ISRNTFPVIIINESITNTQHFINHLEEMTTAYY 119
 DB 60 HGAALPLMLVFGNMTDSFANNG-----NMSBAKRAAFK-LEEMTTAYY 113
 QY 120 YSGIGAVLYAAYIOVSEFCIAAGROIILKROFTHAMROEIGFVDVHVGELNTRLTD 179
 DB 114 YTGIGAVLYAAYIOVSEFCIAAGROIILKROFTHAMROEIGFVDVHVGELNTRLTD 173
 QY 180 DYSKINEGIDRIGMEFHSIATFTGFTGFTGKTLVLITLAI SPVIGISAATWAKILS 239
 DB 174 DYSKINEGIDRIGMEFHSIATFTGFTGFTGKTLVLITLAI SPVIGISAATWAKILS 233
 QY 240 SFDKELLAYAKAGAAVEVLAIRTVAFGOKKELERYKNLEAKGIGIKKAITANI 239
 DB 234 SFDKELLAYAKAGAAVEVLAIRTVAFGOKKELERYKNLEAKGIGIKKAITANI 233
 QY 300 SIGAFLILYASYALAFYGTSLVSEYSIGOVLTVPFESVLGAFSIGASSTIEFAN 359
 DB 294 SMGAFLILYASYALAFYGTSLVSEYSIGOVLTVPFESVLGAFSIGASSTIEFAN 353
 QY 360 ARGAAVEYLFKTIIDNKPISDYSKSGHKPDNIKNGLEFNKVFHVSYSRKEKILKGLINKY 419
 DB 354 ARGAAVEYLFKTIIDNKPISDYSKSGHKPDNIKNGLEFNKVFHVSYSRKEKILKGLINKY 413
 QY 420 QSGQVALVNGSGGKSTTYOIMORLDPDIDGMCIDGQIRITINVRHLEITGVVSOEP 479
 DB 414 KSGQVALVNGSGGKSTTYOIMORLDPDIDGMCIDGQIRITINVRHLEITGVVSOEP 473
 QY 480 VLFATTIENIRYGRENVYMDIEKAVKEANAYDFIMKLPNKEDTLVGERGAOLSGGOKO 539
 DB 474 VLFATTIENIRYGRENVYMDIEKAVKEANAYDFIMKLPNKEDTLVGERGAOLSGGOKO 533
 QY 540 RIATAARLVNRPKILLDEATSAIDPESEAVNOVALDKARKGTTTVIAHRLSTVANNADY 599
 DB 534 RIATAARLVNRPKILLDEATSAIDPESEAVNOVALDKARKGTTTVIAHRLSTVANNADY 593
 QY 600 IAGDDGVIVKGNHDEIMKEKGYFKLYVTMOTRGNEIELEJANATGSKSESALEMSPD 659
 DB 594 IAGDDGVIVKGNHDEIMKEKGYFKLYVTMOTRGNEIELEJANATGSKSESALEMSPD 653
 QY 660 SGSSLLIKRSTRSIAHPOODRKLGTEDLNENPVPSFWIRILKLNSTEMPYPVVGIFC 719
 DB 654 SGSSLLIKRSTRSIAHPOODRKLGTEDLNENPVPSFWIRILKLNSTEMPYPVVGIFC 713
 QY 720 AIINGGLOPAFSIIFSRIGIFRDEDPETKRONSMFVULVLVIGISITFTFLOGTGF 779
 DB 714 AIINGGLOPAFSIIFSRIGIFRDEDPETKRONSMFVULVLVIGISITFTFLOGTGF 773
 QY 780 GRAGELLTRLRKLRVRSMLRODVSWFDPKNTGALTTRLANDAAQYKGAIGSLAVIT 839
 DB 774 GRAGELLTRLRKLRVRSMLRODVSWFDPKNTGALTTRLANDAAQYKGAIGSLAVIT 833
 QY 840 QNTANLGTGIIISLYGWOULTLLLAIVPIALAGVEMKMLSGOALKDKKLEBAGAKIA 899

DB 834 QNTANLGTGIIISLYGWOULTLLLAIVPIALAGVEMKMLSGOALKDKKLEBAGAKIA 893
 QY 900 TEALENFRVYSLTRREKFEYMTAOSLOVPYRNSLRKAHIFGVSESTOAMYPFVACGF 959
 DB 894 TEALENFRVYSLTRREKFEYMTAOSLOIPYRNAMKKAHVFTSFQAMYPFVACGF 953
 QY 960 RFGAYLVANEFNMFODLVFSAIVFGAMAVGVSSFADPAKATVSAHRIITEKPE 1019
 DB 954 RFGAYLVTOQLMFEENVLVFSAIVFGAMAVGVSSFADPAKATVSAHRIITEKPE 1013
 QY 1020 IDSYSPGKLPNTLEONVTNEVFNYPTRPDIPVLOGISLEKKGQTLATVSSGCCSKS 1079
 DB 1014 IDSYSTGKLPNMLEONVOPSGVFNYPTRPSIPVLOGISLEKKGQTLATVSSGCCSKS 1073
 QY 1080 TVVQLERFDPPLAGSVLIDGKEIKHLNQMRAHLGIVSOEPIIFDCSIAENIAYGNS 1139
 DB 1074 TVVQLERFDPPLAGSVLIDGKEIKHLNQMRAHLGIVSOEPIIFDCSIAENIAYGNS 1133
 QY 1140 RVVSHHEIQAQAEANIHFIETLEPKYNTRVGDKGTOLSGGOKORIAIARALYRPHIL 1199
 DB 1134 RVVSYEEIVRAKAEANIHFIETLEPKYNTRVGDKGTOLSGGOKORIAIARALYRPHIL 1193
 QY 1200 LIDEATSAIDTESKYNQVADLDAKAREGRTCIYIAHRLSTIONADLIYVONGKVEHGT 1259
 DB 1194 LIDEATSAIDTESKYNQVADLDAKAREGRTCIYIAHRLSTIONADLIYVONGKVEHGT 1253
 QY 1260 QOLLAOKGYFSMVSVOGAKR 1281
 DB 1254 QOLLAOKGYFSMVSVOGAKR 1275

RESULT 6

us-09-769-097-4
 ; Sequence 4, Application US/09769097
 ; Patent No. US20020055128A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Kimberly Anne Brun
 ; APPLICANT: Richard James Chenery
 ; APPLICANT: Harma Ellens
 ; APPLICANT: John Anthony Feld
 ; APPLICANT: Lin Yue
 ; TITLE OF INVENTION: POLYNUCLEOTIDE AND POLYPEPTIDE SEQUENCES
 ; FILE REFERENCE: GP-50009-C2
 ; CURRENT APPLICATION NUMBER: US/09/769,097
 ; PRIORITY FILING DATE: 2001-01-25
 ; PRIOR APPLICATION NUMBER: 09/208,809
 ; PRIOR FILING DATE: 1998-12-09
 ; PRIOR APPLICATION NUMBER: 09/156,800
 ; PRIOR FILING DATE: 1999-09-17
 ; PRIOR APPLICATION NUMBER: US99/20770
 ; NUMBER OF SEQ ID NOS: 4
 ; SOFTWARE: FastSeq for Windows Version 3.0
 ; SEQ ID NO 4
 ; LENGTH: 1272
 ; TYPE: PRT
 ; ORGANISM: RATTUS RATTUS
 us-09-769-097-4

Query Match 86.7%; Score 5616; DB 10; Length 1272;
 Best Local Similarity 86.3%; Pred. No. 0;
 Matches 1105; Conservative 81; Mismatches 85; Indels 10; Gaps 2;

QY 1 MDEGGRKSAEKNFMKSKKKKKKKPTVSFAMFRKSNMLDRLYMLVGTMAII 60
 DB 1 MELEEDLKGADKNFSKMGKSKK-EKKKKPAVSVLTFMFRAGWLDRLYMLVGTMAII 59
 QY 61 HGAALPLMLVFGNMTDSFANNGISRNTFPVIIINESITNTQHFINHLEEMTTAYY 120
 DB 60 HGAALPLMLVFGNMTDSFANNGISNRNTFPVIIINESITNTQHFINHLEEMTTAYY 110

Qy	121	SGGAGVLAAYIOWSEWCLAAGROILKIRKOPFHAIIMOELGMDYHDGEMTRLTD	180
Dd	111	TGIGAGVLIAAYIOWSTWCLAAGRIHRIKREFHAIIMOELGMDYHVBGLNRLTD	170
Qy	181	VSKINSGIDGXIGMFHSHIAPEFTEFIEIGFIRGKLTUVLAIISPVLGISAIMAKILSS	240
Dd	171	VSKINSGIDGXIGMFQOAMATPFGEIIGFIRGKLTUVLAIISPVLGISAIMAKILSS	230
Qy	241	FTDKELIAYAKAGAAVEVLAIRTVIAFGQKKELERYKNLEEKKGIGIKKALTANIS	300
Dd	231	FTDKELIAYAKAGAAVEVLAIRVIAFGQKKELERYNNLEEKKGIGIKKALTANIS	290
Qy	301	IGAFILIVASYALAAWGTSLVLSSEKSTIGOVLTFFSVLIGAFISIGQASPIEAFAANA	360
Dd	291	MGAAFLIVASYALAAWGTSLVLSIKETTIOQVLTFFSVLIGAFISVQASPIEAFAANA	350
Qy	361	RGAAVEIFKIIIDKNPISIDSYSGKHPDNIKGNLEFKNAHVSIPSRKEXKILGKLNKYO	420
Dd	351	RGAAVEFVSIIDKNPISIDSPSKSGKHPDNIOGNLEFKNIHVSYPSRKDVOLLGKLNLAKE	410
Qy	421	SGQVVALVGNSSCGKSTYVOLMQRXYDPTDQWVCIDGODITITVIRLAREITVVSQEPY	480
Dd	411	SGQVVALVGNSSCGKSTYVOLRLDPLLEGVSLIDGODITITVIRLAREITVVSQEPY	470
Qy	481	LEATTIAENINIGRENTYMDIELEKAYEANAADFIMKLPNKEDTLVGERGAOLSGQOKR	540
Dd	471	LEATTIAENINIGRENTYMDIEIKAYEANAADFIMKLPHKFDTLVGERGAOLSGQOKR	530
Qy	541	IATAIARVNRKIIILDEBATSALDTSSEAVNOVALDKARKGRTIVIAHLSIVNRADVI	600
Dd	531	IATAIARVNRKIIILDEBATSALDTSSEAVNOVALDKARKREGRTIVIAHLSIVNRADVI	590
Qy	601	AAFDGQVIVEKGNDELMAKKEKGIYFKLYMTQTRGNEIELENAEESKSESDELAMSKPOS	660
Dd	591	AAFDGQVIVEKGNDELMBREKGIYFKLYMTQTRAGNEIELENGEACESQDIDNVDMSSKOS	650
Qy	661	GSSLIKRSTRSRJHAPOGDKILGTRKEDLNEWPVPSWRILKINSTEMPYFVAGJFCA	720
Dd	651	GSSLIKRSTRSRJIRGPHDQDELSTREKALDDVNPASWRILKINSTEMPYFVAGJFCA	710
Qy	721	IINGGLOPASTIFSRITIGITREDEPETERKONSNFSLVLEJVLGISITFPLQGTFG	780
Dd	711	IINGGLOPASTIFSRKGVYGTATNDPDEIQROSNLSFLITFGLISITFPLQGTFG	770
Qy	781	KAGEIITKRLRYWFRSRKODVSWFPDPKNTGALTETTLADAQAQVAGALGSLAVITO	840
Dd	771	KAGEIITKRLRYWFRKSMRLRODISWFPDPKNTGALTETTLADAQAQVAGALGSLAVITO	830
Qy	841	NINMLGTGIISLIGMOLTLLLAIPVIAIAGVEMKMLSGOALKDKKELEGAKIAT	900
Dd	831	NINMLGTGIISLIGMOLTLLLAIPVIAIAGVEMKMLSGOALKDKKELEGSGKIAT	890
Qy	901	EAIENFRVYVSLTREQETBYMAOSLOVPPRNSIKRAHI.FGVSESLTOAMWYFSAQCFR	960
Dd	891	EAIENFRVYVSLTREQETBYMAOSLOIPYRNALKRAHFGITFSTOAMWYFSAQCFR	950
Qy	961	FGAVIYAENPEMPODVLVFSATVFGAANAQOVSSFAPDYAKAKVSAAHVIMIEKSPLI	1020
Dd	951	FDALVARELMTFENVLVFSATVFGAANAQOVSSFAPDYAKAKVSAASHIIRLIEKIPEI	1010
Qy	1021	DSVSPHGLKENTLEGNVTENEVENYPTRPDIIVYLOSLISLEKKGOTLALVSSSCCGKST	1080
Dd	1011	DSVSTELKPNMLEGVNAKFNQWMEYPTRPDIIVYLOSLISLEKKGOTLALVSSSCCGKST	1070
Qy	1081	VVOLLEFVYPLAGSVLIDGKEITKHLANQOMLRAHGIVSQEPILTFPQCSIAENIAYGDSNR	1140
Dd	1071	VVOLLEFVYPMAGVIFLKGKEIKOLANQOMLRAHGIVSQEPILTFPQCSIAENIAYGDSNR	1130
Qy	1141	VVSHHEIMQAQKANIHFETLPERKYNTRVGDGTOLSGGQORIAIARALVROPHILL	1200
Dd	1131	VVSHKEIVKAKEANIHOFIDSLPERKYNTRVGDGTOLSGGQORIAIARALVROPHILL	1190
Qy	1201	IDEATSAIDTESEKVVOEALDKARBGRCIYIAHRLSTIQNADLIIVFQNGKVEKHGTIO	1260

[illegible]

```

RESULT 7
US-09-769-097-2
: Sequence 2, Application US/09769097
: Patent No. US20020055128A1
: GENERAL INFORMATION:
: APPLICANT: Kimberly Anne Brun
: APPLICANT: Richard James Chenery
: APPLICANT: Harna Ellens
: APPLICANT: John Anthony Feld
: APPLICANT: Lin Yue
: TITLE OF INVENTION: POLYNUCLEOTIDE AND POLYPEPTIDE SEQUENCES
: TITLE OF INVENTION: ENCODING RAT MDR1A AND SCREENING METHODS THEREOF
: FILE REFERENCE: GP-50009-C2
: CURRENT APPLICATION NUMBER: US/09/769,097
: CURRENT FILING DATE: 2001-01-25
: PRIOR APPLICATION NUMBER: 09/208,809
: PRIOR FILING DATE: 1998-12-09
: PRIOR APPLICATION NUMBER: 09/156,800
: PRIOR FILING DATE: 1999-09-17
: PRIOR APPLICATION NUMBER: US99/20770
: PRIOR FILING DATE: 1999-09-10
: NUMBER OF SEQ ID NOS: 4
: SOFTWARE: FASTSEQ for Windows Version 3.0
: SEQ ID NO 2
: LENGTH: 1272
: TYPE: prt
: ORGANISM: RATTUS RATTUS
US-09-769-097-2

```



```

; TITLE OF INVENTION: Homologue on Chromosome 7p15-21 and Uses Thereof
; FILE REFERENCE: 81994/268611
; CURRENT APPLICATION NUMBER: US/09/873,409
; CURRENT FILING DATE: 2001-06-05
; NUMBER OF SEQ ID NOS: 19
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 5
; LENGTH: 1222
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: No. US20020037522A1e
; LOCATION: (230)..(230)
; OTHER INFORMATION: Xaa at position 230 represents any L amino acid
US-09-873-409-5

```

```

Query Match          54.4%; Score 3525.5; DB 10; Length 1222;
Best Local Similarity 54.1%; Pred. No. 2,2e-262;
Matches 681; Conservative 238; Mismatches 270; Indels 69; Gaps 9;

```

```

QY 51 MLVGTMAIIGGALPLMLVFGNMTDSFANAGISRKPTPVIIINSTNTQHFINHLE 110
DB 1 MIIIGTASLVGACLPMPPLVIGEMSDNLISGCL-----VQNTISFPR--- 44
QY 111 EEMTYATYYSGIGAGVYVAAYIOVSWCLAGROILKIRKOFEMHIMROELGMPDHYV 170
DB 45 -----LTLVYVIGVAAVLIIFGYIOISLIMLTAAKQTKRIKQPFHSHVLADIGMFDSCI 99
QY 171 GELMTPLDVSKEINIGDKIGMFEHSITFEFTGTVGTTRKWKTLVTLAISPLVLGS 230
DB 100 GELMTPRMT-DIDKISDIDGKIALIFQNMSTFSTIGLAVGLVKGMKLTVLVTLSTPLMAS 158
QY 231 AAIKAKLISFTDKELLAYAKAGAAVEVLAIRTVIARGGQKELE----- 277
DB 159 AAACSRWVSIITSKELSAYSKAGAAVEVLSIRTVIARQAQKELOGRFLNITRYAMF 218
QY 278 -----RYKNKLEAKGIGIKKAITANISIGAFLLIYASTALAFWYGTSLV 323
DB 219 YFPQMLISCVLXPKRYQNLKDKMDFQIKKNTASKVSLAAVYFFNMGTGLAFWYGTSLI 278
QY 324 LSSE--YSIGOVLVFSPVLIGAFSIGQASPSIEAFANRGAAYEFKIIDKPSIDYS 381
DB 279 LNEBPYGTIGTVLAVFESSYHSSYCIGAAPHETAIIRGAAPHIFVYIDKPSIDNS 338
QY 382 KSGKRPINIGNLEFKNVHFSYPSREVKILGLNLKVSGQTVALVNSGCGKSTVQL 441
DB 339 TAGKKPESIGTVEFKNVSEFNPSRSIKILKGLNLRKSGFTVALVGLNSGKSTVQL 398
QY 442 MORLYPTDGMVCIIDGODIRTIIVRHLREITGVSEPEVLFATTIAENIRYGRENVYMD 501
DB 399 LQRLYDPDQGFIVDENDIRALVNRHYRKHIGVSOEPVLFTGTTISNNIKYGRDVTDEE 458
QY 502 IEKAVKEANAAYDFIMKLPKFPDLVGERGAQLSGGOKORIAIARALVNRPKILLDEATS 561
DB 459 MEBAAREANAAYDFIMEFPKFMNTLVGERGAQMSGGOKORIAIARALVNRPKILLDEATS 518
QY 562 ALDTESEAVYVALLDKARKGRITTVIAHRLSTVRNADVIAGDGVIVEKNHDELMAKR 621
DB 519 ALDSEKSAVOALEKASGRTTVIAHRLSTIRADLIVTLKQMLAEKGAHAHELMAKR 578
QY 622 GIYFKLVYMQ-TRGNEIELENATGESKSESDLEW-SPKSDSSSLIKRSTRRSIHAPQG 679
DB 579 GLYLSLVMSQDIKKADEQMEHMTYSTERTKNSLPLSHAKSKTSKSDFT----- 624
QY 680 QDRKLGTEDLLENVPVPSFWRLKLNSTEWYFVVGIFCAIINGGLQPARSIFSRNIG 739
DB 625 -DKAESSTQSKESLSEPSVLKILKLNKPEWPFVGLTASVLANGTVAPVPSIIFAITIT 683
QY 740 IFTRPDEPFTKRONSMFSVLVLVGIISFTFELOGTPEBKAGEIILTKRIYVFSMIL 799
DB 684 MF-GNNDKTKTLKHDAEITSMITVILGIVCEVSYFMOGLFETRAGEIILTMRLRHAFKAML 742
QY 800 RODVSKFDDPKMTGALTTRLANDAAQVGAIGSRILAVITTONIANLGTIIISLYQMOL 859

```

```

DB 743 YODIAMPEKEKSTGGTLTTLIAIDTAQIGATGSRIGITLONATNMGLSYISPIYGM 802
QY 860 TLLLIATVPIIAIAGVEMKMLSGOALKDKKELGAGKATFEATEENRTVYSLTRQKE 919
DB 803 TFLIISINPVLAVNGIMETAAITGAFANKDKOELHAGKINTEALENTRTVSLTREKAFE 862
QY 920 YMYAOSLOVPYRNSLRKAHIFGVSEFSITQAMMYPVSACRFAGYIIVANFMFOVLAV 979
DB 863 QMYEEMLOTORHNTSKKQAQIIGSCAFSHAFYFAVYAGRFAGYIILQARMTPEGMFIV 922
QY 960 FSAIVFAMAYGOVYSFAPDPAKAKVSAHVMITEESPLIDYSTSPHGLKPNLTESNVTF 1039
DB 923 FTALVGAAMAIGKTLVLAPEYSKAKSGAAHLEFALKERKPNIDISQEGKKPDKESNLEF 982
QY 1040 NEVVFNPTRPDIIVLOGLSLEVRKQOTLAVVSGSGGKSTVOLLEREPDPLAGSLVD 1099
DB 983 REVSFHYCRPDVFLINGLSISTIRGKTVAFVSSGCGKSTVQLQRLDIPVQGOVLFD 1042
QY 1100 GKELKHLNVOMLRAHLGIVSOEPILPDCSIAENIAYGDNRSVYSHBEIMOAKENIHFF 1159
DB 1043 GVDKKEINVOMLRSQAIIVPOEPLFNCSTIAENIAYGDNRSVYPLDEIKEAANAANHSF 1102
QY 1160 IETLPEKNTFVDPGKQVLSGGOKORIAIARALVNRPHILLDEATSALDTESEKYVQEA 1219
DB 1103 IEGPEKNTNOVGKLGKQVLSGGOKORIAIARALVNRPHILLDEATSALDTESEKYVQEA 1162
QY 1220 LDKAREGRTCIIVAHRLSTIONMDLIVFPNGRVKHEGTHQOLLAQGIYFVSVOA 1277
DB 1163 LDKARTGRTCIVVTHRLSAIONMDLIVLHNGKIKRQGTGHELTRDRDITPKLVANQS 1220

```

```

RESULT 11
US-09-873-409-4
; Sequence 4, Application US/09873409
; Patent No. US20020037522A1
; GENERAL INFORMATION:
; APPLICANT: Sayegh, Mohamed
; TITLE OF INVENTION: A Gene Encoding a Multidrug Resistance Human P-glycoprotein
; FILE OF INVENTION: Homologue on Chromosome 7p15-21 and Uses Thereof
; CURRENT APPLICATION NUMBER: US/09/873,409
; CURRENT FILING DATE: 2001-06-05
; NUMBER OF SEQ ID NOS: 19
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 4
; LENGTH: 1058
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: No. US20020037522A1e
; LOCATION: (66)..(66)
; OTHER INFORMATION: Xaa at position 66 represents any L amino acid
US-09-873-409-4

```

```

Query Match          48.1%; Score 3112.5; DB 10; Length 1058;
Best Local Similarity 56.2%; Pred. No. 9.1e-231;
Matches 601; Conservative 201; Mismatches 220; Indels 47; Gaps 6;

QY 240 SFTDKELLAYAKAGAAVEVLAIRTVIARFGQKKELE----- 277
DB 4 SITSKELSAYSKAGAAVEVLSIRTVIARQAQKELOGRFLNITRYAMFYPQMLISC 63
QY 278 -----RYKNKLEAGIGIKKAITANISIGAFLLIYASTALAFWYGTSLVLSSE--YSI 330
DB 64 VLXEVRYTONLMDADFGIKRTIAKSVSLGAVYFFNMGTVGLAFWYGTSLINDEPGYTI 123
QY 331 GOVLTVFESVLIGAFSIGQASPSIEAFANRGAAYEFKIIDKPSIDYSKSGKRPDI 390
DB 124 GTVLAVFVSVHSSYCIGAAPHETAIIRGAAPHIFVYIDKPSIDNSYIAGIPESTI 183
QY 391 KGNLEFKNVHFSYPSREVKILKGLNLKVSGQTVALVNSGCGKSTVQLDMORLYDPTD 450

```


Db 599 LIDEATSDLTSEGVQALERAEGRTTIVIAHRLSTIKTAHNIIVLVNGKIAEQGH 658
 QY 615 DELMKREKGYFLKVTMQTKGNEIE---LENATGESKESDALEMSPKDGS-----L 664
 Db 659 DELVDKGGAYRKLYEAKRINEQKEADALDADAEDLTNDIAKIKKASASSDLDCKPTT 718
 QY 665 IKRRSTRSIHADQODRKIGTKEDINENVPVPSFRRILK----LNSTWPIFVVGIFCA 720
 Db 719 IDRTGTHKSVSSALSKRP-----PETTPKYSLWTLTKFVASFRRPEIPYMLIGLVFS 771
 QY 721 IINGGLOPASITFSRIGITFREDDEPERK-RONSNMFSLVFLVIGITFSITFEFLOGFTF 779
 Db 772 VLAGGGQPAVLVYAKAISTSLPESQYSKLRHADPFWSLMEFVVGIIQFTQSTGAF 831
 QY 780 GKAGEILTKRLRYVFRSMRLRODVSWFDPKNTGALTTRLANDAAQVKAIGSLAVIT 839
 Db 832 AVCSERILRRARSTAFRTILRODIAFDRKDNSTGALTSLSTETKLSGVSQVTLGITL 891
 QY 840 ONIANIGTGIIISLYGQWLTLLLAIVPIALAGVEMKMLSGQALKDKKEGAGKTA 899
 Db 892 MTSTTLGAALIIIALIGMKLALVCISVYPVLLACGFYREYMLAQFQSRSLAYEGSANA 951
 QY 900 TEALENFRVAVSLTRECKFEYMAQSLQVPRNSLRKAHIFGVFSITQAMTYSAGCF 959
 Db 952 CEATSSIRTVASLTREDVWEIYHAOLDAGRTSLISVNSLSLVAQSQALVEFCVALGF 1011
 QY 960 RFQAYIVANEFMNEQDVLVFSALVFGAMAVGOVSFADPYAKAKYSAHVIMIEKSP 1019
 Db 1012 WYGTTLGHHHEYDIFRFVFCFSEILFGAQSAGVYFSPADPMGRKAKNAAFRRILFRKPO 1071
 QY 1020 IDSSPHGLKPNTEGNAVTFNEVYVNPTRPDIPVLOGLSLEVKKQGTLLVAGSSGCGKS 1079
 Db 1072 IDNMSSEGEKLEYVEGELERRNHFRTPTPEQPVLRGLDLYTKPGQYVALVPSGCGKS 1131
 QY 1080 TVVOLLERFYDPLAGSVLIDKEIKHLNVQMLRAHLGIVSQEPLTDCSTIAENIANGDS 1139
 Db 1132 TTLLEEFYDALINGSLVQCKDISKININSYRSFSLVASEPILYQGTIKENILGIVE 1191
 QY 1140 RVSHHEELMOAKKANIHFFETLPEKYNRVGCKQTQSGGOKORAIARALVROPHIL 1199
 Db 1192 DDVEEELFKCKDANITYDFLMSLPESFNIVVSGKSGMLSGQORVALAKALRDPKIL 1251
 QY 1200 LLDERTSALDTSEKVVQVQALDKAREGRTCIIVIAHRLSTIQNADLIIVFONGKVEHGT 1259
 Db 1252 LLDERTSALDTSEKVVQVQALDAARGRITTAHVRSLTIQKAVIIVFPOGKIVESGTH 1311
 QY 1260 QQLLAQKGIYSWVSQAGAK 1280
 Db 1312 SELVQKGRYYELVNLQSLGK 1332

RESULT 14

; Sequence 2, Application US/09873409
 ; Patent No. US20020037522A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Sayegh, Mohamad
 ; TITLE OF INVENTION: A Gene Encoding a Multidrug Resistance Human P-Glycoprotein
 ; FILE REFERENCE: 81994/268611
 ; CURRENT APPLICATION NUMBER: US/09/873,409
 ; CURRENT FILING DATE: 2001-06-05
 ; NUMBER OF SEQ ID NOS: 19
 ; SOFTWARE: PatentIn version 3.0
 ; SEQ ID NO 2
 ; LENGTH: 812
 ; TYPE: PRT
 ; ORGANISM: Homo sapiens
 ; US-09-873-409-2

Query Match 37.7%; Score 2442; DB 10; Length 812;
 Best Local Similarity 56.8%; Pred. No. 1,9e-179;

Matches 469; Conservative 160; Mismatches 178; Indels 18; Gaps 4;
 QY 455 IDGODIRINRHLREITGVVSQEPVLEFATTAENITVGRGVNTMDIEKAVKANAAYDF 514
 Db 2 VDENDIRALNVRHVRDHHGVVSQEPVLEGTISNNIKYGRDVTDEMEARAARANAAYDF 61
 QY 515 IMKLPNKFDLTVGERGQALSGGOKORIAARALVRNPKILLDDEATSDLTSESAVVOVA 574
 Db 62 IMEPNKNNTLVGERGQALSGGOKORIAARALVRNPKILLDDEATSDLTSESAVVOVA 121
 QY 575 LDKARGRFTTIVIAHRLSTVRNADVIAGFDGVVEKGNDELKKEKGIYFKLVTMQ-TR 633
 Db 122 LEKAKGRFTTIVIAHRLSTIRASDLIVTLKDMGLAEKGAHMLAKKGLVYSILVMSQDIK 181
 QY 634 GNEITLENATGESSESDALEM-SPKDGSSILKRRSTRSIHADQODRKIGTKEDINE 692
 Db 182 KADEMESMTSTETKNTSLPLHSYKTSKSPF-----DKAESSTQCKEI 226
 QY 693 NVPVPSFWRIILKINSTENPVEVVGIFCALINGLOPAFSITFSKILGIFTRDEDEPERKQ 752
 Db 227 SLPEVSLKIKLTKNPEWPFVVLGTLASVNGTVHPVFSIIFAKIITMF-GNNDKTYLKH 285
 QY 753 NSNMFVLEFLVIGITFSITFEFLOGFTPEKAGEILTKRLRYVFRSMRLRODVSWFDPDKNT 812
 Db 286 DAEIYSMIFVLIVGICFVSFEMQGLFGRAGELLMLRHLRAFKAMLYODIANTWDERENS 345
 QY 813 TGALETRLANDAAQVKAIGSLAVITQNIANTGIIISLYGQWLTLLLAIVPIAI 872
 Db 346 TGGITTLIALIDIAQIQAGATGSRIGVLTQNTVMGLSVISIFYGEMFELIAPVLAV 405
 QY 873 AGVENEKMLSGQALKDKKELEGAKIATEIENFTVYSLTRECKFEYMAQSLQVPRN 932
 Db 406 TGMIEATANTGPRANDKDELKHAIGIATEALENITFTVSLTREKFEYMEQEMBLDQHRN 465
 QY 933 SLRKAHIEGVFSITQAMVFSYAGCFRQAYIVANEFMNEQDVLVFSALVFGAMAVGO 992
 Db 466 TSKRAQIDISCTAFSAPAFYFAAGFRGAVLIQAGNTPRCMNVITFALAYGMAIGK 525
 QY 993 VVSFAPDVAKAVYSAHVIMIEKSPIDISYSPHGLKPNTEGNAVTFNEVYVNPTRPDI 1052
 Db 526 TVLVAEYSKAKSGAHLFALLEKRPNDISRSQEEKKDPDTCGNLEFREFVSFFYCRPDV 585
 QY 1053 PVIQGISLEVKKGOTIALVSGSGCKSTVYVOLLERFYDPLAGSVLIDKEIKHLNVQMLR 1112
 Db 586 FILRGLSLIEGKTVAFVSGSGCKSTSVOLLRYDVOQOVLFDVDAKELNVQMLR 645
 QY 1113 AHLGIYSOPELFDCSIENIAYGDSRVVSHHEILMOAKKANIHFFETLPEKYNTRVG 1172
 Db 646 SQIAIYPOEPVLEFNCISIAENIAYGDSRVVPLEIKKAANMANIHSFIEGLPEKYNTOVG 705
 QY 1173 DKGTQSGGOKORIAARALVROPHILLDDEATSDLTSEKVVQVQALDKAREGRTCI 1232
 Db 706 LKGAQLSGGOKORIAARALRLOKPKILLDDEATSDLTSEKVVQVQALDAARGRITCLAV 765
 QY 1233 AHLSTIQNADLIIVFONGKVEHGTQOOLAKOGIYFMSVQAA 1277
 Db 766 THRLSATIONADLIIVLHNGKIKHQGHQELLRNDIYFKLVNAQS 810

RESULT 15

; US-10-101-388-3
 ; Sequence 3, Application US/10101388
 ; Patent No. US20020162142A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Johal, Gurmukh S
 ; APPLICANT: Multani, Dilbag P
 ; APPLICANT: Briggs, Steven P
 ; TITLE OF INVENTION: GENES AND METHODS FOR MANIPULATION OF GROWTH
 ; FILE REFERENCE: 5718-81 (035718/205794)
 ; CURRENT APPLICATION NUMBER: US/10/101,388
 ; CURRENT FILING DATE: 2002-03-19
 ; PRIOR APPLICATION NUMBER: 09/711,562
 ; PRIOR FILING DATE: 2000-11-13

```

; PRIOR APPLICATION NUMBER: 60/164,886
; PRIOR FILING DATE: 1999-11-12
; NUMBER OF SEQ ID NOS: 3
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 3
; LENGTH: 1394
; TYPE: PRN
; ORGANISM: Zea mays
US-10-101-388-3

```

```

Query Match      30.1% Score 1950; DB 9; Length 1394;
Best Local Similarity 36.1% Pred. No. 2,6e-141;
Matches 479; Conservative 235; Mismatches 468; Indels 146; Gaps 28;

```

```

QY 25 NEKKERKPTSTFAMFRYSNMIDRLVMTLGTWAAITIHGAALPLMLVLCGNMTDSFANAIGI 84
   | : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 113 NDSKKPTPPAALRDLRFRAADGLDCAIMLIGTLGALVHCISLPVLRFFADLVDSFGSHAD 172
   | : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 85 SRKKTFPVLIINSTNNTOHFINLHEEMTYAAYYSIGAGVLVAAYIOVSEWCLAAGR 144
   | : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 173 DDEPTMVRLLVVK-----YAFYFLVVGAAIWAASSWAEISCMWTGER 212
   | : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 145 QLIKIRKOFFHAIMROEIGMFDVHDVGLNTRLDDVSKINEIGDXTGMFFSHIAIFFT 204
   | : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 213 QSTRMIRIRLDALRQDVSFFDT-----DVRASDVITAIADAVVVOGR--HQPETGQP 264
   | : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 205 GTVGTGRTGKWLTVILAIISPYLISAIAIMAKILSFTDKELLAYAKAGAV-----AEE 258
   | : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 265 HPLHGLHNG-----RLRGVHGRVAAG-----AGHARGAANRRHRAER 304
   | : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 259 VLAALRTVIANG-----GQKKELEYKNLBEAK-----GIGIKKAITANT 299
   | : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 305 RRAARQALVPOQRALGROHRHGAAGARADTDAGVRMGARGAGLLGGAGGADRLRQL 364
   | : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 300 SIGAA-----FLIYAYALAFWYGTSLVLSSEYSIGOVLFVFSVLGARSIGQASP 352
   | : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 365 RQGARPRRHLLRHLLLRPALVRRPRARAHORRRARHNTM---FSVMIGG-GPRQSA 420
   | : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 353 STEAFANRGAAYELEFKTIIDKPSIDSYSGKHPDNKGNLEFKVNFESYPSRKEVYKIL 412
   | : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 421 SMAAPAKARVAAKIFRITIDHRPGIS--SRDGAPEBSVTGRVEMRGVDFAYPSRPDPIL 478
   | : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 413 KGLNLKYVSGQTVALVGNSSGCKSTVQLMORLYDPTDGVWCIDGODIRTIANRHLREIT 472
   | : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 479 RGFSLSVAGRTIALVSSSGSKSTVSLIERFYDPAGQILLDGHLRSLRLMLRQI 538
   | : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 473 GVYSGEPVLEATTIAENIRYGR--ENVTMDIEIKAVKEANAYDFIKLDPNKPEDTVGERS 530
   | : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 539 GLVSGEALFATSIRENILLGRDSQATLAEEMEAARVANAHSFIITKLDPGYDTVOGERS 598
   | : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 531 AOLSGGOKORAIARALVRNPKIILDEATSAIDTESAVVOVALDKARKGRTT-VIAH 589
   | : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 599 LQLSGGOKORAIARAKMLKPKPAILLDEATSAIDSESEKLYOEAIDRFMGKRTLLCDRAT 658
   | : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 590 RLSTVRNADYIAGFDVIVIEKGNHDEIMK--EKGIYFKLYMTMOTRGNIELENAATGESK 647
   | : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 659 GCPSPAKADVAVYVLOGGAVSEMHADEIMAKGNGTYAKILIRMOEQAHEALVNA----- 713
   | : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 648 SEDALMSPKDSSSLIKRRS-----TRR-----STHAPOGDRKLGTR 688
   | : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 714 RRSARPSSARNSVSPIMTRNSYGRSPYSRLSDFTSLIHDPNHHRTMADKQ 773
   | : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 689 DLNENVPVPSFMRILKNSTEMPEYFVVGICAIINGLOPAFSLIFSRLIGTFTDEDEPE 748
   | : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 774 -LAFRAGASSFLRLARNNSPEMAYALAGSIGVCCSFSATFAVTLISAVLSVYA--PD 829
   | : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 749 TKRQNSNMFVLFVLVGIISFTTF--LOGFTGKAGELITLRLRYMFRSMLRQDVSMF 806
   | : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 830 PRYMKREIAYCYLLIGMSSAALLENTVOHVFWMDTVGEGULTRKRVREKMFVAFVNEIAMF 889
   | : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 807 DDEKNTTGALTTRLANDAQAQVGAIGSRILAVITTONIANIGITIIISLYGQQLILLAI 866
   | : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 890 DADENASAKVTATLADDAQONRSALIGRISIVLQNSALMLVACTAGFVLQWRALVLIAY 949
   | : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :

```

```

QY 867 VPIIAIAGVE---MKMLSGQALKDKKELEG---AGIATEAIEINFRTVVSILTREQKFE 919
   | : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 950 FPLVGAATVLOKKFMKGFSG-----DLEAHARATQIAGAVANLRTVAFAFAERKIT 1002
   | : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 920 YMYAOSLOVPRNSLRKAHIFGVSFISIQAMMYFESTYACCFRGAIVLANETMNFQDVLV 979
   | : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 1003 GLEFANLRQPLRRCFWKKGQIAGSGYGAOFLLYASALGLMYAAMLVKGHVSDFSRITRY 1062
   | : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 980 FSAIYEGAMAQVSSFAPDYAKAKVSAAHVIMITEKSPILDSYSPGLK----PNTLEG 1035
   | : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 1063 FMYLWVSANGAAETILNAPDFIKGRANRSVFETIDRKTVEY---PHDYAAPVDPDGA 1119
   | : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 1036 NTFENEVNNTTRPDIPVLOGLSLEVRKGGOTLALVSSGGCKSTVOLLERFYDPLAGS 1095
   | : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 1120 KYELKHVDLPSRPDIQVFRDLSRABARQTLALVGPSSGCKSSVLALVQRFYKPTSGR 1179
   | : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 1096 VLDGKEIKHLNVMWLRNHLGIVSOEPIIFDCSTAEINTAYAGDNSRVVSHETIMOAKKAN 1155
   | : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 1180 VLLDGKDVAKYMLRRLKRVAVAVPOEPLFAASIHENIAYGREG--ATEAEVEEAAAOAN 1237
   | : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 1156 IHFTETLPEKYNTRVGDGKTQLSGGCKORAIARALVROPHILLDDEATSAIDTSEKXV 1215
   | : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 1238 AHRFTALPEGYRTQVGERGVQLSGGQRORTAIRALVKKQALVLDDEATSAIDAESEK 1297
   | : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 1216 VQEAIDKAREGETCIVIAHRLSTIQNADLIYFQNGKVKHEGHQOULLAO--KGIYFSNV 1273
   | : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 1298 VQEAIDKAREGETCIVIAHRLSTIQNADLIYFQNGKVKHEGHQOULLAO--KGIYFSNV 1273
   | : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 1274 SVOAGAKR 1281
   | : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 1358 QLAADGR 1365
   | : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :

```

Search completed: December 9, 2002, 17:12:06
Job time : 151 secs

GenCore version 5.1.3
Copyright (c) 1993 - 2002 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: December 9, 2002, 16:20:53 ; Search time 639 seconds
(without alignments)
192.720 Million cell updates/sec

Title: US-09-672-725c-2
Perfect score: 6477
Sequence: 1 MPPEGKSGSAEKNFWKMK.....LAAQKGIYSVSVQAGAKR 1281

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283224 seqs, 96134422 residues

Total number of hits satisfying chosen parameters: 283224

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database :
1: p1r1:*
2: p1r2:*
3: p1r3:*
4: p1r4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB ID	Description
1	5861	90.5	1280	1 DVHU1	multidrug resist
2	5651.5	87.3	1276	1 DVHY1C	multidrug resist
3	5633	87.0	1276	2 A34786	multidrug resist
4	5296	81.8	1276	1 DVMS1	multidrug resist
5	5226.5	80.7	1277	2 JH0502	multidrug resist
6	4990	77.0	1104	1 DVMS1A	multidrug resist
7	4905.5	75.7	1279	1 DVHU3	multidrug resist
8	4856	74.4	1276	1 DVMS2	multidrug resist
9	4818.5	74.4	1281	2 I48123	multidrug resist
10	4803	68.3	1287	2 S53692	multidrug resist
11	4422	68.3	1287	2 S53692	multidrug resist
12	3209.5	49.6	1321	2 T42228	multidrug resist
13	3205.5	49.5	1321	2 T42228	multidrug resist
14	2857	44.1	1294	2 T19982	multidrug resist
15	2819	43.5	1289	2 D77889	multidrug resist
16	2818.5	43.5	1275	2 T31073	multidrug resist
17	2741	42.3	1321	2 T23476	multidrug resist
18	2739	42.3	1321	2 S27337	multidrug resist
19	2674.5	41.3	655	1 DVHY2C	multidrug resist
20	2651	40.9	1283	3 A47377	multidrug resist
21	2602.5	40.2	1286	2 T02187	multidrug resist
22	2596	40.1	1292	2 T48007	multidrug resist
23	2589.5	40.0	1278	2 E86155	multidrug resist
24	2564.5	39.6	1302	2 A41249	multidrug resist
25	2521	38.9	1229	2 D85023	multidrug resist
26	2518	38.9	1229	2 T52319	multidrug resist
27	2518	38.9	1230	2 E85023	multidrug resist
28	2496.5	38.5	1302	2 B41249	multidrug resist
29	2473	38.2	1229	2 F86155	multidrug resist

30	2441	37.7	1408	2 T43261	multidrug resist
31	2437.5	37.6	1333	2 H85202	multidrug resist
32	2431	37.5	1286	2 A42150	multidrug resist
33	2423	37.4	1310	2 S30328	multidrug resist
34	2406	37.1	1266	2 T22090	multidrug resist
35	2398	37.0	1268	2 T22094	multidrug resist
36	2380	36.7	1302	2 S30327	multidrug resist
37	2345	36.2	1254	2 S27328	multidrug resist
38	2330.5	36.0	1222	2 T14805	multidrug resist
39	2307	35.6	1233	2 T04251	multidrug resist
40	2297.5	35.5	1245	2 G86404	multidrug resist
41	2289	35.3	1254	2 T30855	multidrug resist
42	2287	35.3	1307	2 T30882	multidrug resist
43	2277	35.2	1362	2 T41534	multidrug resist
44	2269.5	35.0	1318	2 T21266	multidrug resist
45	2246	34.7	1247	2 F86405	multidrug resist

ALIGNMENTS

RESULT 1
DVHU1 multidrug resistance protein 1 - human
N:Alternate names: P-glycoprotein 1
C:Species: Homo sapiens (man)
C>Date: 31-Dec-1990 #sequence_revision 18-Aug-1995 #text_change 19-Jan-2001
C:Accession: A34914; PS0162; S15500; A25059; S43838; I52238; I65204
R:Chen, C.; Clark, D.; Ueda, K.; Pastan, I.; Gottesman, M.M.; Roninson, I.B.
J. Biol. Chem. 265, 506-514, 1990
A:Title: Genomic organization of the human multidrug resistance (MDR1) gene and origin
A:Reference number: A34914; MUID:90094448; PMID:1967175
A:Accession: A34914
A:Molecule type: DNA
A:Residues: 1-1280 <CHE>
A:Cross-references: GB:M29447; GB:J05168; NID:q187496; PIDN:AAAS59576.1; PID:g386862
R:Kloka, N.; Yamano, Y.; Komano, T.; Ueda, K.
submitted to the EMBL data library, April 1991
A:Description: Transcriptional regulation of multidrug resistance gene (MDR1) express
A:Reference number: S15500
A:Accession: S15500
A:Molecule type: DNA
A:Residues: 1-22, 'R' <RI2>
A:Cross-references: EMBL:X58723; NID:g34522; PIDN:CAAA1558.1; PID:g34523
R:Chen, C.; Chin, J.E.; Ueda, K.; Clark, D.P.; Pastan, I.; Gottesman, M.M.; Roninson, I.B.
Cell 47, 381-389, 1986
A:Title: Internal duplication and homology with bacterial transport proteins in the m
A:Reference number: A25059; MUID:87028230; PMID:2876781
A:Accession: A25059
A:Molecule type: mRNA
A:Residues: 1-184, 'V', 186-1280 <CH2>
A:Cross-references: GB:M14758; NID:g187468; PIDN:AAAS59575.1; PID:g307180
R:Chambers, T.C.; Pohl, J.; Glass, D.B.; Kuo, J.F.
Biochem. J. 299, 309-315, 1994
A:Title: Phosphorylation by protein kinase C and cyclic AMP-dependent protein kinase
A:Reference number: S43838; MUID:94220047; PMID:7909431
A:Accession: S43838
A:Molecule type: protein
A:Residues: 656-689 <CHA>
R:Geckeler, V.; Weger, S.; Probst, H.
Biochem. Biophys. Res. Commun. 169, 796-802, 1990
A:Title: mdrl/P-glycoprotein gene segments analyzed from various human leukemic cell
A:Reference number: I52238; MUID:90290529; PMID:1972623
A:Accession: I52238
A:Status: translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 178-215 <RES>

A:Cross-references: GB:M37724; NID:g183537; PIDN:AAA8047.1; PID:g553314
A:Accession: 165204
A:Status: translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 800-856 <RR2>
A:Cross-references: GB:M37725; NID:g183538; PIDN:AAA8048.1; PID:g553315
A:Comment: This is an integral membrane protein overproduced in multidrug-resistant cell
structurally and functionally unrelated lipophilic anticancer drugs.
C:Genetics:
A:Gene: GDB:PGY1; MDR1
A:Cross-references: GDB:120712; OMIM:171050
A:Map position: 7q21-7q21
A:Superfamily: multidrug resistance protein; ATP-binding cassette homology
C:Keywords: ATP; duplication; glycoprotein; nucleotide binding; P-loop; phosphoprotein;
F:1-638,653-1280/Region: duplication
F:49-350/Domain: hydrophobic <HB1>
F:351-637/Domain: hydrophobic <HB1>
F:410-604/Domain: ATP-binding cassette homology <ABC1>
F:427-434/Region: nucleotide-binding motif A (P-loop)
F:551-555/Region: nucleotide-binding motif B
F:638-708/Domain: linker <LIN>
F:709-993/Domain: hydrophobic <HB2>
F:994-1280/Domain: hydrophobic <HL2>
F:1053-1249/Domain: ATP-binding cassette homology <ABC2>
F:1070-1077/Region: nucleotide-binding motif A (P-loop)
F:1196-1200/Region: nucleotide-binding motif B
F:91-94,99/Binding site: carbohydrate (Asn) (covalent) #status predicted
F:433/Binding site: ATP (Lys) #status predicted
F:661,667,671/Binding site: phosphate (Ser) (covalent) (by CAMP-dependent kinase) #statu
F:1076/Binding site: ATP (Lys) #status predicted

Query Match 90.5%; Score 5861; DB 1; Length 1280;
Best Local Similarity 90.7%; Pred. No. 0;
Matches 1163; Conservative 55; Mismatches 60; Indels 4; Gaps 4;

```

OY      1  MPEGGKRGSA- EKNFMWKKSKKNEKKKPYSTAMRYSNWLDRLMLVGTMAAI 59
DB      1  MDEGGKRGCAKKKNEFKLNKSEK-DKEKKPYSVSMRYSNWLKLMVGTIAAI 59
OY      60  IHCAALPLMLVGNMTDSFANAGISRNTFPVAINESITNTQHFINHLEEMTYAVY 119
DB      60  IHGAGLPMLLVGEMTDIFANAG-NLEDMSINIRSDINDTGFPM-LEEDMTRYAVY 117
OY      120 YSGIGAGVLAAYIQVSEWCLAGROILKIROFFAHIMRODIGEDVDHDELTRTLD 179
DB      120 YSGIGAGVLAAYIQVSEWCLAGROILKIROFFAHIMRODIGEDVDHDELTRTLD 177
OY      180 DVSKINGIGDCKIGMFHSAIFPFTGFIYFTGRGKMLTLVLAISPVLAGLSAAIWA 239
DB      178 DVSKINGIGDCKIGMFHSAIFPFTGFIYFTGRGKMLTLVLAISPVLAGLSAAVAK 237
OY      240 SFTDKRELLAYAKAGAAVEVLAIRTVIAFGGKKLEKYNKNLEBAKCIIGIKKAITANI 299
DB      238 SFTDKRELLAYAKAGAAVEVLAIRTVIAFGGKKLEKYNKNLEBAKCIIGIKKAITANI 297
OY      300 SIGAFLITIASTALAWYGTSTVLSSEYSGOVLVFPFSVLGAFSISQASPSIEAFAN 359
DB      298 SIGAFLITIASTALAWYGTSTVLSSEYSGOVLVFPFSVLGAFSISQASPSIEAFAN 357
OY      360 ARGAAVEIFKIIDNKPISIDSYSGKGRPDNIKGNLEKYNHESYPSRKEVKILGINKLV 419
DB      358 ARGAAVEIFKIIDNKPISIDSYSGKGRPDNIKGNLEKYNHESYPSRKEVKILGINKLV 417
OY      420 QSGQVALVNGSGCKSTTVQLMKRLDPTDGNVICDGDIFRTINVRHLREITGVNSQEP 479
DB      418 QSGQVALVNGSGCKSTTVQLMKRLDPTDGNVSVGDQIRITINVRFLREIIGVNSQEP 477
OY      480 VLEFATTIENIRYGRENTYMDIEKAVKEANAADYFIKMLPNKFDTLVGERGALSGGQXQ 539
DB      478 VLEFATTIENIRYGRENTYMDIEKAVKEANAADYFIKMLPNKFDTLVGERGALSGGQXQ 537
OY      540 RIAIARALVRNPKILLDEATSAIDTSEANVOVALDKARKGRITTVIAHRLSTVRNADY 599

```

```

DB      538 RIAIARALVRNPKILLDEATSAIDTSEANVOVALDKARKGRITTVIAHRLSTVRNADY 597
OY      600 IAFEDGVIYEKGNHDELKKEKGIYFKLTVMOTRGNEIELENAGSESDALAMSPKO 659
DB      598 IAFEDGVIYEKGNHDELKKEKGIYFKLTVMOTAGNEVELEMADESKSIDALEMSND 657
OY      660 SGGSLIKRSTRSIAPOGOKRIGTKEDLNENPVPSPFIRILKINSTENPYPVVGIFC 719
DB      658 SRSILIKRSTRSIVSGSQADRKLTKEALDESIPIPSFPMIMKLNLTENPYPVVGIFC 717
OY      720 AIIINGGQPAFSTIFSRIGITFRDDEPETKRONSNMVSFLVJGIIISFTIFLOGTFE 779
DB      718 AIIINGGQPAFSTIFSRIGITFRDDEPETKRONSNMVSFLVJGIIISFTIFLOGTFE 777
OY      780 GKAGEIITKRLRIVWFSMLRQDVSWFDDPKNTGALTTRLANDAAQVGAIGSRNAVIT 839
DB      778 GKAGEIITKRLRIVWFSMLRQDVSWFDDPKNTGALTTRLANDAAQVGAIGSRNAVIT 837
OY      840 QNIANLGGIISLIYGMQLTLILAIVPIIAIAGVVEKMKLSGQALKDKKELEGAKIA 899
DB      838 QNIANLGGIISLIYGMQLTLILAIVPIIAIAGVVEKMKLSGQALKDKKELEGAKIA 897
OY      900 TEATENEFYVSLTRBEQFEYMTAQSLQVYRYSIKRAHIFGVSFSTQAMVYFVAGCF 959
DB      898 TEATENEFYVSLTRBEQFEYMTAQSLQVYRYSIKRAHIFGVSFSTQAMVYFVAGCF 957
OY      960 RFGATLVANEMNODVILVFSATVFGAMAVGVQVSPFADYAKAKVSAHVIMIEKSP 1019
DB      958 RFGATLVANEMNODVILVFSATVFGAMAVGVQVSPFADYAKAKVSAHVIMIEKSP 1017
OY      1020 IDSYPHGLKPNTEGNTFENVEVFNTPRPDIPVLOGLSLEFKKQGTALVSSGCGKS 1079
DB      1018 IDSYPHGLKPNTEGNTFENVEVFNTPRPDIPVLOGLSLEFKKQGTALVSSGCGKS 1077
OY      1080 TYVOLLERFPYPLAGSVLIDGKEIKHLVQMLRAHLIVSGEPIFLPDCSIAENIAYGDS 1139
DB      1078 TYVOLLERFPYPLAGSVLIDGKEIKHLVQMLRAHLIVSGEPIFLPDCSIAENIAYGDS 1137
OY      1140 RVVSHHEIQAQKAEINHHFETLEPEKYNTRYGDKGTOLSGGQORAIARALVROPHIL 1199
DB      1138 RVVSHHEIQAQKAEINHHFETLEPEKYNTRYGDKGTOLSGGQORAIARALVROPHIL 1197
OY      1200 LIDEATSAIDTSESEKVOBALDKAREGRTCIYIAHRLSTIONADLIYVFGNGKREHGT 1259
DB      1198 LIDEATSAIDTSESEKVOBALDKAREGRTCIYIAHRLSTIONADLIYVFGNGKREHGT 1257
OY      1260 QQLLAQGIYFSWVSVOAGAKR 1281
DB      1258 QQLLAQGIYFSWVSVOAGAKR 1279

```

RESULT 2
DVH1C
Multidrug resistance protein 1 - Chinese hamster
N:Alternate names: P-glycoprotein pgp1
C:Species: Cricetus griseus (Chinese hamster)
C:Date: 31-Dec-1990 #sequence revision 30-Sep-1992 #text change 19-Jan-2001
C:Accession: A38696; C38696; B38696; A27126; S33768; I52823
R:Devine, S.E.; Hussain, A.; Davie, J.P.; Melzer, P.W.
J. Biol. Chem. 266, 4545-4555, 1991
A:Title: Full length and alternatively spliced pgp-1 transcripts in multidrug-resistant
A:Reference number: A38696; MUID:91154265; PMID:1671863
A:Accession: A38696
A:Molecule type: mRNA
A:Residues: 1-1276 <DEV>
A:Cross-references: GB:M59253; NID:g191154; PIDN:AAA37004.1; PID:g191155
A:Accession: C38696
A:Molecule type: mRNA
A:Residues: 108-1276 <DB1>
A:Cross-references: GB:M59254; NID:g191156; PIDN:AAA37005.1; PID:g191157
A:Experimental source: clone ADX185
A:Accession: B38696

A:Molecule type: mRNA
A:Residues: 1-32,771-1276 <DE2>
A:Cross-references: GB:M59252; NID:g191152; PIDN:AAA37003.1; PID:g191153
A:Experimental source: clone ADX124
R:Endicott, J.A.; Juranka, P.F.; Sarangi, F.; Gerlach, J.H.; Deuchars, K.L.; Ling, V.
Mol. Cell. Biol. 7, 4075-4081, 1987
A:Title: Simultaneous expression of two P-glycoprotein genes in drug-sensitive Chinese H
A:Accession: A27126
A:Status: not compared with conceptual translation
A:Molecule type: mRNA
A:Residues: 706-1276 <END>
A:Cross-references: GB:M17897; NID:g191158; PIDN:AAA37006.1; PID:g191159
R:Zastawny, R.L.; Ling, V.
Biochim. Biophys. Acta 1173, 303-313, 1993
A:Title: Structural and functional analysis of 5' flanking and intron 1 sequences of the
A:Accession: S33768
A:Status: translation not shown
A:Molecule type: DNA
A:Residues: 1-21 <ZAS>
A:Cross-references: EMBL:L03286
R:Teeter, L.D.; Eckersberg, T.; Tsai, Y.; Kuo, M.T.
Cell Growth Differ. 2, 429-437, 1991
A:Title: Analysis of the Chinese hamster P-glycoprotein/multidrug resistance gene pgp1
A:Accession: I52823
A:Status: translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-21 <RES>
A:Cross-references: GB:S81975; NID:g240862
C:Comment: This is an integral membrane protein overproduced in multidrug-resistant cell
structurally and functionally unrelated lipophilic antitumor drugs.
A:Gene: pgp1
C:Superfamily: multidrug resistance protein; ATP-binding cassette homology
C:Keywords: alternative splicing; ATP; duplication; glycoprotein; nucleotide binding; P-
F:407-601/Domain: ATP-binding cassette homology <ABC1>
F:424-431/Region: nucleotide-binding motif A (P-loop)
F:548-552/Region: nucleotide-binding motif B
F:1050-1246/Domain: ATP-binding cassette homology <ABC2>
F:1067-1074/Region: nucleotide-binding motif A (P-loop)
F:1193-1197/Region: nucleotide-binding motif B
F:87, 91, 96/Binding site: carbonylate (asn) (covalent) #status predicted
F:1073/Binding site: ATP (Lys) #status predicted
Query Match 87.3% Score 5651.5; DB 1; Length 1276;
Best Local Similarity 86.8%; Pred. No. 0;
Matches 113; Conservative 77; Mismatches 85; Indels 7; Gaps 3;
QY 1 MPEGGRGSAEKNEKWKSKKNEKKEKPPVSTFAMFRYNNMIDRLYMVGMAAI 60
DB 1 MFEEDFSGRDKNFKMKGRSKK-EKKEKKPVSVFTMRAYAGHRLMVGTLAAII 59
QY 61 HGAALPLMLVFGNMTDFANAG-ISKNTFFVITINESTNNTQFINHLEEMTVAYY 119
DB 60 HGVALPLMLVFGNMTDFASVGNITPNAI-----NNAQVNASDIFGLEEMTVAYY 114
QY 120 YSGIGACVLAAYIOVSFWCLAGROILKIRKOFHAIKROEIGMFDVHDGELNTRITD 179
DB 115 YTGIGACVLAAYIOVSFWCLAGROILKIRKOFHAIKROEIGMFDVHDGELNTRITD 174
QY 180 DVSFKINGIDGKIGMPFHSIATFTFGYVGRGKMLTVLILAIISPVGLSAAIMAKILS 239
DB 175 DVSFKINGIDGKIGMPFHSIATFTFGYVGRGKMLTVLILAIISPVGLSAAIMAKILS 234
QY 240 SFDDKELAAKAGAAVEVLAIRTVLAIFGQKKELEERYKNLEAKGIGIKATAITNT 299
DB 235 SFDDKELAAKAGAAVEVLAIRTVLAIFGQKKELEERYKNLEAKGIGIKATAITNT 294
QY 300 SIGAAFLILYASVLAFAWYGTSLVLSSEYSGQVTVFFSVLIGAFSIGQASPSIEAFAN 359
I:|||||

DB 295 SMGAFLILYASVLAFAWYGTSLVLSSEYSGQVTVFFSVLIGAFSIGQASPSIEAFAN 354
QY 360 ARGAAVEIFKIIDNKPSIDSYSKSGHKPNIKNGNEFKVNVHSYSSREKVIILGLNKKV 419
DB 355 ARGAAVEIFKIIDNKPSIDSYSKSGHKPNIKNGNEFKVNVHSYSSREKVIILGLNKKV 414
QY 420 QSGQVALVNSCGSGSTTVQMLQRLYPTDGMVCIQDQIRTIINVRRLREITGVSOEP 479
DB 415 QSGQVALVNSCGSGSTTVQMLQRLYPTDGMVCIQDQIRTIINVRRLREITGVSOEP 474
QY 480 VLFATTAENIRYGRNVMTDEIEKAVKANAYDITMLPKKFPPTIVGKERAOISGQKO 539
DB 475 VLFATTAENIRYGRNVMTDEIEKAVKANAYDITMLPKKFPPTIVGKERAOISGQKO 534
QY 540 RIIAIAIYRNKILLDLDEATSLDTESEAVVOVALDKARKGRTTIVIAHRLSTVRNADY 599
DB 535 RIIAIAIYRNKILLDLDEATSLDTESEAVVOVALDKARKGRTTIVIAHRLSTVRNADY 594
QY 600 IAGFDGVIIEKGNHDELMKEKGIYFKLVMTQTRNGNEIEENATGESKSDALEKSPD 659
DB 595 IAGFDGVIIEKGNHDELMKEKGIYFKLVMTQTRNGNEIEENATGESKSDALEKSPD 654
QY 660 SSSSLIKRSTPRSTHAQODRKLGKEDINENVPVPSFWRLIKLNSTEPYFVYVIGRC 719
DB 655 SSSSLIKRSTPRSTHAQODRKLGKEDINENVPVPSFWRLIKLNSTEPYFVYVIGRC 714
QY 720 AIIINGLOPAFSIIFSRIGIFTRDEDEPETKRONSNFSLFVLCIGISFTYFFLOGTFE 779
DB 715 AIIINGLOPAFSIIFSRIGIFTRDEDEPETKRONSNFSLFVLCIGISFTYFFLOGTFE 774
QY 780 GKAGSILKRLRYVFNFSMLRDVSWFDDPKNTGALTTLRLANDAQQVGAIGSLRAYIT 839
DB 775 GKAGSILKRLRYVFNFSMLRDVSWFDDPKNTGALTTLRLANDAQQVGAIGSLRAYIT 834
QY 840 QNIANLGTGLIISLYQMOLTLILAIIVPIATAGVVEKMLSGQALDKKLEBAGKTA 899
DB 835 QNIANLGTGLIISLYQMOLTLILAIIVPIATAGVVEKMLSGQALDKKLEBAGKTA 894
QY 900 TEALENRIVVSLTRDQKFEYVYAOGLQVPRYNSLRKAHIFGVSPSLIQAMMYFSGCF 959
DB 895 TEALENRIVVSLTRDQKFEYVYAOGLQVPRYNSLRKAHIFGVSPSLIQAMMYFSGCF 954
QY 960 RFQAYILVAFNFMFOVLLVFSAYVGAAGVSSFPADYAKAVSAHYIMITEKSPIL 1019
DB 955 RFQAYILVAFNFMFOVLLVFSAYVGAAGVSSFPADYAKAVSAHYIMITEKSPIL 1014
QY 1020 IDSYSPHGLKPNLTGEGVTFNEVFNYPTRPDIPYLQGLSLTEKKGOTLALVSSGCGKS 1079
DB 1015 IDSYSPHGLKPNLTGEGVTFNEVFNYPTRPDIPYLQGLSLTEKKGOTLALVSSGCGKS 1074
QY 1080 TVQVLLERFYPDLASVYLLIDGKEIKHLVWQRLAHGLGVSOEPLIFDCAIENIAYGDS 1139
DB 1075 TVQVLLERFYPDLASVYLLIDGKEIKHLVWQRLAHGLGVSOEPLIFDCAIENIAYGDS 1134
QY 1140 RYVSHIELIQAAKKANIHFTETLPEKYNIVGDKGTQLSGGOKORAIARALVAPRPHL 1199
DB 1135 RYVSHIELIQAAKKANIHFTETLPEKYNIVGDKGTQLSGGOKORAIARALVAPRPHL 1194
QY 1200 LIDEATSLDTESEKVVQVQALDKARKGRTTIVIAHRLSTQANDLIVVONGKVEKHGTH 1259
DB 1195 LIDEATSLDTESEKVVQVQALDKARKGRTTIVIAHRLSTQANDLIVVONGKVEKHGTH 1254
QY 1260 QOLLAQKGIYSVMSVQAGAKR 1281
DB 1255 QOLLAQKGIYSVMSVQAGAKR 1276
RESULT 3
A34786
multidrug resistance protein 1a - mouse
C:Species: Mus musculus (house mouse)
C:Date: 13-Jul-1990 #sequence_revision 13-Jul-1990 #text_change 02-Feb-2001
C:Accession: A34786; A35671

Query Match	87.08:	Score 5633:	DB 2:	Length 1276:
Best Local Similarity	87.08:	Pred. No. 0:		
Matches 1115:	Conservative 74:	Mismatches 85:	Indels 8:	Gaps 4:
QY 1	MOPEGKRSASAKNFWMKMGKSKSKNEKKERKPTVSTFAMFRISNMLDRYMLHVGCMALI	60		
DB 1	MELEEDLGRADNMSTSKMGKSKK-EKKERKPPAVSYLIRFAGWGLDRYMLHVGGLIAI	59		
QY 61	HGAAFLPLMLVGNMTDSEFANAG-ISRNKTEPPVLINESTINNTQHIINHLEEMTVAAY	119		
DB 60	HGVALPLMLLIGDMTDSFASVGNKSNST-----NNSEADKRAMPAK-LEEMTVAAY	113		
QY 120	YSGIGAGVLAANYIOVSEFWCLAAGROILIKTKQFFHAIMROELGMPDVGDELMTRLTD	179		
DB 114	YTGIGAGVLAANYIOVSEFWCLAAGROILIKTKQFFHAIMROELGMDVDHDELMTRLTD	173		
QY 180	DVSKINGIDCKIGMFHSHIATFFTGFTVGTFRGMKLTLLVLAISPYLGISAIAKNTILS	239		
DB 174	DVSKINGIDCKIGMFEOAMTFPGFGITIGTRBMKLTLLVLAISPYLGISAIAKNTILS	233		
QY 240	SFTDKELLAAKAGAAVEVLAIRTVIAFGGOKKELEERYKNLDEAKIGIKRAITANI	299		
DB 234	SFTKELHAATKAGAAVEVLAIRTVIAFGGOKKELEERYNNLDEAKIGIKRAITANI	293		
QY 300	SIGAFLLLIYASVALAFWGTSLVLSSEYSIGQVLTVEFSVLICAFSIOGASSIAFAN	359		
DB 294	SMGAFLLLIYASVALAFWGTSLVLSKEYSIGQVLTVEFSVLICAFSIOGASSIAFAN	353		
QY 360	ARGAAVELEFKIIDKRPISIDSYSKSGHKPDNIGNLEFNHVSYPSEKREKILGLNTKY	419		
DB 354	ARGAAVELEFKIIDKRPISIDSEKSGHKPDNIOGNLEFNHVSYPSEKREKILGLNTKY	413		
QY 420	OSGQVALVNGSGCGKSTTYVOLMORLYDPTDGNVCIDGODITTIWRLRLREITVGSQEP	479		
DB 414	KSGQVALVNGSGCGKSTTYVOLMORLYDPTDGNVSLIDGODINTINWRLRLREITVGSQEP	473		
QY 480	VLFATIIAENIRYGRHENVYTMDEIERAKVREANAYDTIMKLPNKFPTLVGEGRGAOLSGGOK	539		
DB 474	VLFATIIAENIRYGRHENVYTMDEIERAKVREANAYDTIMKLPNKFPTLVGEGRGAOLSGGOK	533		
QY 540	RAIARALVNRNKKILLDEEATSAIDTESEAVNOVALDKARKGRTTVIAHRLSTVRNADV	599		
DB 534	RAIARALVNRNKKILLDEEATSAIDTESEAVNOVALDKARKGRTTVIAHRLSTVRNADV	593		

[illegible]

Query Match	81.8%;	Score 5296;	DB 1;	Length 1276;
Best Local Similarity	80.3%;	Pred. No. 1.4e-299;		
Matches 1029;	Conservative 130;	Mismatches 115;	Indels 8;	Gaps 5

QY 660 SCSLLKRRSTRSHIAPGCGDRKLTGKTDLENVPEVSWRLKLNSTEMPEYFVVGIFC 719

D6 657 SSSPLL-RRSYRSYVRHAKDQERRLSMKCAVDVEDPLVSWFRLMLNSTEMPYLLVGVLC 715

QY 720 AIIINGGLQAFASTIESRIRIGIFTRDEDPETKRONMSVLEVLGIIISLTFELQSTF 779

D6 716 AVINGICIVFVAVISRIIVGVFSRDDHETTKQONLTSFLFWGLISFVLYFEGGTFE 775

QY 780 GRAGELLIRKLRVWFPRSKLRODVSWFDDPKNTGALTTRILANDAAQVKAIGSRLAITY 839

D6 776 GRAGELLIRKRVRYWPKSMLRODISWFDHKNSTSTLTTRILASPDASSVKAMGARLAVYT 835

QY 840 QNIANIIGTCIIISLTYGQMTLLILAIVPIAIAAGVWEMKMLSGQALKDKKKELEGAKIA 899

D6 836 QNVANLIGCVILSLTYGQMLLLVLIPLVYLGGIIEMKLLSQALKDKKKELEISGIA 895

QY 900 TEAIEFRTVSLREOKFEYMAOSLQVPRNSLRKAHIGVSFSITQAAMVFSYACGF 959

D6 896 TEAIEFRTVSLTREOKFEYMAOSLQVPRNMMKKAHVIGITFSITQAAMVFSYACF 955

QY 960 RFGALVAVNEMNRQDVLVFSALVFGAMAVGVSRPADYAKAKVASAAHYIMIEKSPL 1019

D6 956 RFGALVAVQDLMTETENWMLVSAVFGAMAGNTSSFPADYAKAKVASASHIRIIEKTFPE 1015

QY 1020 IDYSFPHGLNTEGNTFENEVFENYPTRPDIIVLQGLSLEYKGGOTLALVSSGGGKS 1079

D6 1016 IDYSTEGILKPTLLEGNNKENGVOFNPTRPINPVYLGSLSEYKGGOTLALVSSGGGKS 1075

QY 1080 TVVQGLEFYPPLVGSVLIDGKEIKHLNVQMLRAHLGIVSDEPTLFDSCIAENIAYGDNS 1139

D6 1076 TVVQGLEFYPDPMAGSVFLDGKEIKQLNVQMLRAHLGIVSDEPTLFDSCIAENIAYGDNS 1135

QY 1140 RVVSHETIMQAQKANTHHELETPLEKYNTRVGDKGTQLSGGQKORAIARALVRPHLL 1199

D6 1136 RAVSHETIVRAKAKANIHQIDSLPEKYNTRVGDGKGTQLSGGQKORAIARALVRPHLL 1195

QY 1200 LLDATGALTDTESKRYVQVQLDKARBGRTCIYIAHRLSTIONADLLIVFONGKVKSHGTH 1259

D6 1196 LLDATGALTDTESKRYVQVQLDKARBGRTCIYIAHRLSTIONADLLIVFONGKVKSHGTH 1255

QY 1260 QQLLAQKGIYFSMWVSQVAGAKR 1281

D6 1256 QQLLAQKGIYFSM-VQAGAKR 1275

RESULT 5

JH0502

P-glycoprotein - rat

N:Alternate names: multidrug resistance protein mdr1b

C:Species: Rattus norvegicus (Norway rat)

C.Date: 31-Mar-1992 #sequence,revision 31-Mar-1992 #text_change 02-Feb-2001

C.Accession: JH0502; S22353; S22352

R:Silveman, J.A.; Raunio, H.; Gant, T.W.; Thorgirsson, S.S.

Gene 106, 229-236, 1991

A:Title: Cloning and characterization of a member of the rat multidrug resistance

A:Reference number: JH0502; MUID:92039081; PMID:1682220

A:Accession: JH0502

A:Molecule type: mRNA

A:Residues: 1-1277 <STL>

A:Cross-references: EMBL:X61104; NID:956890; PIDN:CAA43416.1; PID:91334219

A:Accession: S22352

A:Status: Preliminary

A:Molecule type: DNA

A:Residues: 1212-1226, '1', 1228-1270, 'SV', 1271-1277 <DE2>

A:Cross-references: EMBL:X61103; NID:956888; PIDN:CAA43415.1; PID:91334218

A:Cross-references: EMBL:X61103; NID:956888; PIDN:CAA43415.1; PID:91334218

QY 299 ISGAAFLIYAAALAFWGTSTVLSSEKSTQVITVFESVILGAFSTGASPSIEFA 358
 C:Species: Homo sapiens (man)
 C:Date: 31-Dec-1990 #sequence_revision 31-Dec-1990 #text_change 19-Jan-2001
 Db 121 ISMGAAFLIYAAALAFWGTSTVLSSEKSTQVITVFESVILGAFSTGASPSIEFA 180
 QY 359 NARGAAVEIFKIDNKRPSIDYSKSGHKNPKINGNLEFNKVFHSPSPREKVVILKGLTK 418
 Db 181 NARGAAVEEVIKIDNKRPSIDYSKSGHKNPKINGNLEFNKVFHSPSPREKVVILKGLTK 240
 QY 419 VOSQOTVALVNGSCGCGSTVQOLMORLYDPDGMWCIDGDIRTINVRHLREITGVVSOE 478
 Db 241 VKSQOTVALVNGSCGCGSTVQOLMORLYDPDGMVWSIDGDIRTINVRHLREITGVVSOE 300
 QY 479 PVLEPATTIAENIRYRGRENVTMDIEFAKVEAANAADPIMKLPKPEITLVEGGAOLSGGOK 538
 Db 301 PVLEPATTIAENIRYGRDVTMDIEFAKVEAANAADPIMKLPKPEITLVEGGAOLSGGOK 360
 QY 539 ORIAIAALVRNKKILLDEATSAIDTSEAAVQALDKARKGRTIYIAHRLSTVNAD 598
 Db 361 ORIAIAALVRNKKILLDEATSAIDTSEAAVQALDKARKGRTIYIAHRLSTVNAD 420
 QY 599 VIAGFDGVIVEGNDELMEKKEGFEFKLVMTQGTNGNETELENANGESKSDALEMSPPK 638
 Db 421 VIAGFDGVIVEGNDELMEKKEGFEFKLVMTQGTNGNETELENANGESKSDIDLMDSK 480
 QY 659 DSGSSILKRSTRSIHAQODRKLGTKEDELNENVPVSEFWRLIKLNSTEMPVFGIF 718
 Db 481 DSGSSILKRSTRSIICGPHDDRSLSTEALDEDPVSPASFWRLIKLNSTEMPVFGIF 540
 QY 719 CAIINGGLPAPSAIISRIIGITFDEDEPERKROSNMSEVLFVIGITISFTIFFLOGFT 778
 Db 541 CAIINGGLPAPSAIISRIIGITFDEDEPERKROSNMSEVLFVIGITISFTIFFLOGFT 600
 QY 779 EKGAGELIKRLRIYWFERSMLKODVSNFDPKNTGALTTRLANDAAQVGAIGSLAVI 838
 Db 601 EKGAGELIKRLRIYWFERSMLKODVSNFDPKNTGALTTRLANDAAQVGAIGSLAVI 660
 QY 839 TONIAMGTGIIISITLYGMOLTLILLALVPIITAIAGVEMKMSGQALKRKELEAGXI 898
 Db 661 FOMIAMLGIIISITLYGMOLTLILLALVPIITAIAGVEMKMSGQALKRKELEAGXI 720
 QY 899 ATPEATENFTVSLTRECKFEYMTAQSLOVPEYRNSLRKAHIFGVSITQAMMYFSYAC 958
 Db 721 ATPEATENFTVSLTRECKFEYMTAQSLOVPEYRNSLRKAHIFGVSITQAMMYFSYAC 780
 QY 959 FRFGAIVLANEFMNPQDVLVEFAIVFGAAMVGVSSFADYAKAKVSAHYMITKEXP 1018
 Db 781 FRFGAIVYQOLMPENVLLVFSALVFGAAMVGVSSFADYAKAKVSAHYMITKEXP 840
 QY 1019 LIDTSPHGLKRNTELEGNVTENEVFNYPTRPDIPLVQGLSLEVKKGQTLATVSGSGCK 1078
 Db 841 EIDTSPHGLKRNTELEGNVTENEVFNYPTRPDIPLVQGLSLEVKKGQTLATVSGSGCK 900
 QY 1079 STVVQVLLERFPDPLAGSVLIDGKEIKHLNOMLRAHLGIYSQEPILFDGSIENIATYAGN 1138
 Db 901 STVVQVLLERFPDPLAGSVLIDGKEIKHLNOMLRAHLGIYSQEPILFDGSIENIATYAGN 960
 QY 1139 SNVVSHEIIMQAAKEMNIHFIETLPEKYNTRVGDGKGTOLSGGOKRIATAALVROPHI 1198
 Db 961 SNVVSHEIIMQAAKEMNIHFIETLPEKYNTRVGDGKGTOLSGGOKRIATAALVROPHI 1020
 QY 1199 LILDEATSAIDTSEKVVQALDKARKGRTIYIAHRLSTIONADLIYVQNGKVEHGT 1258
 Db 1021 LILDEATSAIDTSEKVVQALDKARKGRTIYIAHRLSTIONADLIYVQNGKVEHGT 1080
 QY 1259 HQQLLAQGIYFSVSVQAGAKR 1281
 Db 1081 HQQLLAQGIYFSVSVQAGAKR 1103

RESULT 7
 VHVH3
 multilidrug resistance protein 3 - human

N:Alternate names: P-glycoprotein MDR3
 C:Species: Homo sapiens (man)
 C:Date: 31-Dec-1990 #sequence_revision 31-Dec-1990 #text_change 19-Jan-2001
 Db 121 ISMGAAFLIYAAALAFWGTSTVLSSEKSTQVITVFESVILGAFSTGASPSIEFA 180
 QY 359 NARGAAVEIFKIDNKRPSIDYSKSGHKNPKINGNLEFNKVFHSPSPREKVVILKGLTK 418
 Db 181 NARGAAVEEVIKIDNKRPSIDYSKSGHKNPKINGNLEFNKVFHSPSPREKVVILKGLTK 240
 QY 419 VOSQOTVALVNGSCGCGSTVQOLMORLYDPDGMWCIDGDIRTINVRHLREITGVVSOE 478
 Db 241 VKSQOTVALVNGSCGCGSTVQOLMORLYDPDGMVWSIDGDIRTINVRHLREITGVVSOE 300
 QY 479 PVLEPATTIAENIRYRGRENVTMDIEFAKVEAANAADPIMKLPKPEITLVEGGAOLSGGOK 538
 Db 301 PVLEPATTIAENIRYGRDVTMDIEFAKVEAANAADPIMKLPKPEITLVEGGAOLSGGOK 360
 QY 539 ORIAIAALVRNKKILLDEATSAIDTSEAAVQALDKARKGRTIYIAHRLSTVNAD 598
 Db 361 ORIAIAALVRNKKILLDEATSAIDTSEAAVQALDKARKGRTIYIAHRLSTVNAD 420
 QY 599 VIAGFDGVIVEGNDELMEKKEGFEFKLVMTQGTNGNETELENANGESKSDALEMSPPK 638
 Db 421 VIAGFDGVIVEGNDELMEKKEGFEFKLVMTQGTNGNETELENANGESKSDIDLMDSK 480
 QY 659 DSGSSILKRSTRSIHAQODRKLGTKEDELNENVPVSEFWRLIKLNSTEMPVFGIF 718
 Db 481 DSGSSILKRSTRSIICGPHDDRSLSTEALDEDPVSPASFWRLIKLNSTEMPVFGIF 540
 QY 719 CAIINGGLPAPSAIISRIIGITFDEDEPERKROSNMSEVLFVIGITISFTIFFLOGFT 778
 Db 541 CAIINGGLPAPSAIISRIIGITFDEDEPERKROSNMSEVLFVIGITISFTIFFLOGFT 600
 QY 779 EKGAGELIKRLRIYWFERSMLKODVSNFDPKNTGALTTRLANDAAQVGAIGSLAVI 838
 Db 601 EKGAGELIKRLRIYWFERSMLKODVSNFDPKNTGALTTRLANDAAQVGAIGSLAVI 660
 QY 839 TONIAMGTGIIISITLYGMOLTLILLALVPIITAIAGVEMKMSGQALKRKELEAGXI 898
 Db 661 FOMIAMLGIIISITLYGMOLTLILLALVPIITAIAGVEMKMSGQALKRKELEAGXI 720
 QY 899 ATPEATENFTVSLTRECKFEYMTAQSLOVPEYRNSLRKAHIFGVSITQAMMYFSYAC 958
 Db 721 ATPEATENFTVSLTRECKFEYMTAQSLOVPEYRNSLRKAHIFGVSITQAMMYFSYAC 780
 QY 959 FRFGAIVLANEFMNPQDVLVEFAIVFGAAMVGVSSFADYAKAKVSAHYMITKEXP 1018
 Db 781 FRFGAIVYQOLMPENVLLVFSALVFGAAMVGVSSFADYAKAKVSAHYMITKEXP 840
 QY 1019 LIDTSPHGLKRNTELEGNVTENEVFNYPTRPDIPLVQGLSLEVKKGQTLATVSGSGCK 1078
 Db 841 EIDTSPHGLKRNTELEGNVTENEVFNYPTRPDIPLVQGLSLEVKKGQTLATVSGSGCK 900
 QY 1079 STVVQVLLERFPDPLAGSVLIDGKEIKHLNOMLRAHLGIYSQEPILFDGSIENIATYAGN 1138
 Db 901 STVVQVLLERFPDPLAGSVLIDGKEIKHLNOMLRAHLGIYSQEPILFDGSIENIATYAGN 960
 QY 1139 SNVVSHEIIMQAAKEMNIHFIETLPEKYNTRVGDGKGTOLSGGOKRIATAALVROPHI 1198
 Db 961 SNVVSHEIIMQAAKEMNIHFIETLPEKYNTRVGDGKGTOLSGGOKRIATAALVROPHI 1020
 QY 1199 LILDEATSAIDTSEKVVQALDKARKGRTIYIAHRLSTIONADLIYVQNGKVEHGT 1258
 Db 1021 LILDEATSAIDTSEKVVQALDKARKGRTIYIAHRLSTIONADLIYVQNGKVEHGT 1080
 QY 1259 HQQLLAQGIYFSVSVQAGAKR 1281
 Db 1081 HQQLLAQGIYFSVSVQAGAKR 1103

Query Match 75.7%; Score 4905.5; DB 1; Length 1279;
 Best Local Similarity 75.5%; Pred. No. 75-277;
 Matches 973; Conservative 131; Mismatches 165; Indels 19; Gaps 9;

QY 1 MPEGRGK-----SAEKNFMWKKSKKKEKKEK-PTVSTFAMPRRSNMDRLYLIV 53
 Db 1 MLEAAKNGTAMRPTSAEGDF-ELGISSKQKRRKTKVKNMIGVLTFRYSMDQKLFMSL 59
 QY 54 GTMAIIGHAALPLMLVFGNMVTSFANAGISRNKRPPTVINSINNNOHFINHLEEM 113
 Db 60 GTMAIAGSGPLMLMIVFGEMTDKFDVT--AGNESPVNFSLSLNPK-----IIEEM 113
 QY 114 TTYAYYSGIAGVLYAAIYVSWFLCAAGROLIKRKQFPAHIMQEIIGWFDVHVGEL 173
 Db 114 TTYAYYSGIAGVLYAAIYVSWFLCAAGROLIKRKQFPAHIMQEIIGWFDVHVGEL 173

RESULT 7
DVH03
multidrug resistance protein 3 - human

DMS2 8
 multidrug resistance protein 2 - mouse
 N:Alternate names: P-glycoprotein MDR2
 C:Species: Mus musculus (house mouse)
 C:Date: 31-Dec-1990 #sequence-revision 31-Dec-1990 #text-change 19-Jan-2001
 C:Accession: A30409; S70711
 R:GROS, P.; Raymond, M.; Bell, J.; Housman, D.
 Mol. Cell. Biol. 8, 2770-2778, 1988
 A:Title: Cloning and characterization of a second member of the mouse mdr gene family
 A:Reference number: A30409; MUID:88302195; PMID:3405218
 A:Accession: A30409
 A:Molecule type: mRNA
 A:Residues: 1-1276 <NS>
 A:CROSS-references: GB:J03398; NID:g199109; PIDN:AAA39516.1; PID:g387428
 R:Kirschner, L.S.
 Nucleic Acids Res. 24, 2829-2834, 1996
 A:Title: De novo generation of simple sequence during gene amplification.
 A:Reference number: S70711; MUID:96313253; PMID:8759018
 A:Accession: S70711
 A:Status: nucleic acid sequence not shown; translation not shown
 A:Molecule type: DNA
 A:Residues: 43-92 <KIN>
 A:CROSS-references: EMBL:U46839; NID:g1228142; PIDN:AC52722.1; PID:g1228143
 A:Note: The nucleotide sequence was submitted to the EMBL Data Library, January 1996
 C:Comment: This is an integral membrane protein overproduced in multidrug-resistant C
 structurally and functionally unrelated lipophilic antitumor drugs.
 C:Genetics:
 A:Gene: mdr2
 C:Superfamily: multidrug resistance protein; ATP-binding cassette homology
 C:Keywords: ATP; duplication; glycoprotein; nucleotide binding; P-loop; transmembrane
 F:1-637/653-1276/Region: duplication
 F:409-603/Domain: ATP-binding cassette homology <ABC1>
 F:426-433/Region: nucleotide-binding motif A (P-loop)
 F:550-554/Region: nucleotide-binding motif B
 F:1049-1245/Domain: ATP-binding cassette homology <ABC2>
 F:1067-1074/Region: nucleotide-binding motif A (P-loop)
 F:1192-1196/Region: nucleotide-binding motif B
 F:68,94/Binding site: carbohydrate (Asn) (covalent) #status predicted
 F:432/Binding site: ATP (Lys) #status predicted
 F:1072/Binding site: ATP (Lys) #status predicted


```

OY 357 FANAGAAVEIFKTIIDNKPISIDSYKSGHKPDNIKGNLKNVHFSPSKREVKILKGN 416
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 354 FANAGAAVYFDDIIDNNPKIDSSEKGHKPDNIKGNLKNVHFSPSKRANIKILKGN 413
OY 417 LKVGSGQTVALVNSGGCKSTTVOLMORLPTDGMWICIGODIRITNVRHRLREITGVY 476
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 414 LKVGSGQTVALVNSGGCKSTTVOLMORLPTDGMWICIGODIRITNVRHRLREITGVY 473
OY 477 QEPVLFATTAENIRYRGENVYDETEKAVKEANAYDFIMKLPNKFDVLVGERGAQLSGC 536
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 474 QEPVLFATTAENIRYRGENVYDETEKAVKEANAYDFIMKLPNKFDVLVGERGAQLSGC 533
OY 537 OKORIAIARLVNPKIILDEATSAIDTSESEVVOVALDKARKGKGTIVIAHRLSTVN 596
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 534 OKORIAIARLVNPKIILDEATSAIDTSESEVVOVALDKARKGKGTIVIAHRLSTVN 593
OY 597 ADVIAGPDGVIEKGNHDELMKEKGYIKRYLVTMTGRNEILENATGSKSESDALEMS 656
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 594 ADVIAGPDGVIEKGNHDELMKEKGYIKRYLVTMTGRNEILENATGSKSESDALEMS 653
OY 657 PKDSGSSLIKRRSTRSIIHAPOGODRKLGTKEDLNENVPVSEFMILKLNSTEMPYVYG 716
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 654 P-NGWAKARIPRSTKSKLSPH-QNRLDETEWELDANVPVSEFMILKLNSTEMPYVYG 711
OY 717 IFCATIIINGIOPAFSIIISRIIGITFRDEDEPTKRONSMFVFLVIGIISFTFFLOG 776
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 712 IFCATIIINGIOPAFSIIISRIIGITFRDEDEPTKRONSMFVFLVIGIISFTFFLOG 770
OY 777 FFEKGAGELITRLRYMVRSMRLRODVSMEDPKRTATLTRLANDAAYKGAIGSRLA 836
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 771 FFEKGAGELITRLRYMVRSMRLRODVSMEDPKRTATLTRLANDAAYKGAIGSRLA 830
OY 837 VITONIANIGIITSLIYGMOLTLILAIVPIAIAGVEMKMLSGOALKDKKELEAG 896
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 831 LAQMTANIGIITSLIYGMOLTLILAIVPIAIAGVEMKMLSGOALKDKKELEAG 890
OY 897 KIATAIENFTVSLTPEOKPEVYVYASLOVPRNSIKRAHIFGVSEFTTQAMMTFSYA 956
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 891 KIATAIENFTVSLTPEOKPEVYVYASLOVPRNSIKRAHIFGVSEFTTQAMMTFSYA 950
OY 957 GCFRGAIVLANEFNFODVLVFSALVFGAMAVGVSSFAPADYKAKKVAANVIMILK 1016
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 951 GCFRGAIVLANEFNFODVLVFSALVFGAMAVGVSSFAPADYKAKKVAANVIMILK 1010
OY 1017 SPLIDSYSPHGLKPNLTLEGNTEVENVNPTRPDIPVLQGLSLEVKKGOTLALVSSGC 1076
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1011 SPLIDSYSPHGLKPNLTLEGNTEVENVNPTRPDIPVLQGLSLEVKKGOTLALVSSGC 1070
OY 1077 GKSTVOLLERFYDPLAGSVLIDGKEIKHLNVMOLRAHIGIVSOEPLIDCSIAENIANG 1136
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1071 GKSTVOLLERFYDPLAGSVLIDGKEIKHLNVMOLRAHIGIVSOEPLIDCSIAENIANG 1130
OY 1137 DNSRVYSHETMOAKRANIHHEITLPEKYNTRVGDKTQOLSGGOKORIAIARLVNRP 1196
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1131 DNSRVYSHETMOAKRANIHHEITLPEKYNTRVGDKTQOLSGGOKORIAIARLVNRP 1190
OY 1197 HILLDEATSAIDTSESEVVOVALDKARKGKGTIVIAHRLSTVNADLIYVQNKVKEH 1256
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1191 HILLDEATSAIDTSESEVVOVALDKARKGKGTIVIAHRLSTVNADLIYVQNKVKEH 1250
OY 1257 GTHOQIILAQGIYFSMVSQVQAK 1280
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1251 GTHOQIILAQGIYFSMVSQVQAK 1274

```

RESULT 9
148123

P-glycoprotein isoform III - Chinese hamster

C:Species: Cricetulus griseus (Chinese hamster)

C:Date: 02-Jul-1996 #sequence_revision 02-Jul-1996 #text_change 02-Feb-2001

C:Accession: 148123

R:Endicott, J.A.; Sarangl, F.; Ling, V.

```

DNA Seq. 2, 89-101, 1991
A:Title: Complete cDNA sequences encoding the Chinese hamster P-glycoprotein gene fam
A:Reference number: 148121; M0ID:92135896; PMID:1685679
A:Accession: 148123
A>Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: mRNA
A:Residues: 1-1281 <RES>
A:Cross-references: GB:60042; NID:919168; PIDN:AA68885.1; PID:919169
C:Superfamily: multidrug resistance protein; ATP-binding cassette homology
C:Keywords: ATP; glycoprotein; nucleotide binding; P-loop
F:412-606/Domain: ATP-binding cassette homology <ABC1>
F:429-436/Region: nucleotide-binding motif A (P-loop)
F:1054-1250/Domain: ATP-binding cassette homology <ABC2>
F:1071-1078/Region: nucleotide-binding motif A (P-loop)

Query Match          74.4%; Score 4818.5; DB 2; Length 1281;
Best Local Similarity 73.6%; Pred. No. 7.9e-272;
Matches 948; Conservative 147; Mismatches 176; Indels 17; Gaps 8;

OY 1 MDEPGGRKSGSAEK-----NFMKMKKKKKKKKKKKKKPT-VSFFAMFRYSNMILRLMLVYG 54
    || ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1 MDEAARNGTARPRGTVGEDFELGSLNSGKRNKKKKNVILGPTLFRSDWQDKLEMLDG 60
OY 55 TMAATIHGAALPLMLVFGNMNDSFA-NAGISRNKTPVITINESTITNNQOFIHMLEEM 113
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 61 TMAIAHAGSLPLMLVFEGMTDKFVNAG--NFSLPVNFSLSMINPGR----ILEEM 113
OY 114 TTYAAYYSGIGAGVLAAYIOVSFWCLAGROILKIRQFHAIRKQFEDVHVGEL 173
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 114 TTYAAYYSGIGAGVLAAYIOVSFWCLAGROILKIRQFHAIRKQFEDVHVGEL 173
OY 174 NTRLDDVSKINEIGIGKIGMEFHSIATFTGVTGFRGKTLVILATISPVYIGSAI 233
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 174 NTRLDDVSKINEIGIGKIGMEFHSIATFTGVTGFRGKTLVILATISPVYIGSAI 233
OY 234 WAKIISFTDEKLAYAGAAVEYLAITVTYAFEGCKELEERYKNLEAKGIGIKR 293
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 234 WAKIISFTDEKLAYAGAAVEYLAITVTYAFEGCKELEERYKNLEAKGIGIKR 293
OY 294 AITANISGAIFLLIYASALAFWYGTSLVSEYSIGQVLYTFEFSYLIGAFSIGASPS 353
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 294 AITANISGAIFLLIYASALAFWYGTSLVSEYSIGQVLYTFEFSYLIGAFSIGASPS 353
OY 354 IEAFANAGAAVEIFKTIIDNKPISIDSYKSGHKPDNIKGNLKNVHFSPSKREVKILK 413
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 354 IEAFANAGAAVEIFKTIIDNKPISIDSYKSGHKPDNIKGNLKNVHFSPSKREVKILK 413
OY 414 GLNLKVGSGQTVALVNSGGCKSTTVOLMORLPTDGMWICIGODIRITNVRHRLREITG 473
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 414 GLNLKVGSGQTVALVNSGGCKSTTVOLMORLPTDGMWICIGODIRITNVRHRLREITG 473
OY 474 VVSOEPLVFTTAENIRYRGENVYDETEKAVKEANAYDFIMKLPNKFDVLVGERGAQL 533
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 474 VVSOEPLVFTTAENIRYRGENVYDETEKAVKEANAYDFIMKLPNKFDVLVGERGAQL 533
OY 534 SGGOKORIAIARLVNPKIILDEATSAIDTSESEVVOVALDKARKGKGTIVIAHRLST 593
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 534 SGGOKORIAIARLVNPKIILDEATSAIDTSESEVVOVALDKARKGKGTIVIAHRLST 593
OY 594 VRNADVIAGDDGVIEKGNHDELMKEKGYIKRYLVTMTGRNEILENATGSKSESDAL 653
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 594 VRNADVIAGDDGVIEKGNHDELMKEKGYIKRYLVTMTGRNEILENATGSKSESDAL 653
OY 654 EMSPKDSGSSLIKRRSTRSIIHAPOGODRKLGTKEDLNENVPVSEFMILKLNSTEMPY 712
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 654 GMP-NGWAKARIPRSTKSKLSPH-QNRLDETEWELDANVPVSEFMILKLNSTEMPY 712
OY 713 FVVGICATIIINGIOPAFSIIISRIIGITFRDEDEPTKRONSMFVFLVIGIISFTFF 772
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 713 FVVGICATIIINGIOPAFSIIISRIIGITFRDEDEPTKRONSMFVFLVIGIISFTFF 771
OY 773 FLOGFTKAGELITRLRYMVRSMRLRODVSMEDPKRTATLTRLANDAAYKGAIG 832
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

```

```

Db 772 FLOGFTFGKAGELLITRLSRMAFKAMLRQDMSPDDYKNSGALSTRLATDRAQVOGATG 831
QY 833 SRLAVITQNTANTGCTGIISLIYGMQTLTLLLAIVPIIAIAGVEMKMLSGOALKDKREL 892
Db 832 TRALLIQAONTANTGCTGIISLIYGMQTLTLLLSVPFIASGIVEMKMLAGNAKDKREL 891
QY 893 EGACKIATEAIENFRVYSLTRROKFEYMAOSLOVPYNSLRKAHIFGVSTIQAMMY 952
Db 892 EAACKIATEAIENFRVYSLTRROKFEYMAOSLOVPYNSLRKAHIFGVSTIQAMMY 951
QY 953 FSYAGCRRFAGYIVANFNFDVLVFSALVIGAMAVGVSSFPADYAKAVSAAHVIM 1012
Db 952 FSYAGCRRFAGYIVANFNFDVLVFSALVIGAMAVGVSSFPADYAKAVSAAHVIM 1011
QY 1013 IIEKSPILDSYSPHGLKPNLTGCVTFNEVYVFPTRPDIPYLOGISLEVKKGOTIALVG 1072
Db 1012 LFEROPILDSYSGEGLPMDKFEVSTFNEVYVFPTRPANPVLQGLSLEVKKGOTIALVG 1071
QY 1073 SSGCGKSTVVOGLLREYDPLAGSVLIDGKEIKHLNVQMLRAHLGIYSOEPILEDCSIAEN 1132
Db 1072 SSGCGKSTVVOGLLREYDPLAGSVLIDGKEIKHLNVQMLRAHLGIYSOEPILEDCSIAEN 1131
QY 1133 IAYGNSRVVSHHEIMQAKKANIHHEITLPEKYNTRVQDKGTQLSGCGKORITAIARAL 1192
Db 1132 IAYGNSRVVSHHEIMQAKKANIHHEITLPEKYNTRVQDKGTQLSGCGKORITAIARAL 1191
QY 1193 VQOPHILLDEATSLDTESEKVOEALDKAREGRTCIIVAHRLSTIONADLIYVPONGK 1252
Db 1192 IKOPRYLLDEATSLDTESEKVOEALDKAREGRTCIIVAHRLSTIONADLIYVPONGK 1251
QY 1253 VKEHGTHQOLLAKQGIYFSMVSVOAGAK 1280
Db 1252 VKEHGTHQOLLAKQGIYFSMVSVOAGAK 1279

```

RESULT 10

```

S41646
P-glycoprotein - rat
C:Species: Rattus norvegicus (Norway rat)
C:Date: 13-Jan-1995 #sequence_revision 13-Jan-1995 #text_change 02-Feb-2001
C:Accession: S41646; S22354
R:Brown, P.C.; Thorgerirsson, S.S.; Siljeverman, J.A.
Nucleic Acids Res. 21, 3885-3891, 1993
A:Title: Cloning and regulation of the rat mdr2 gene.
A:Reference number: S41646; MUID:93376516; PMID:8103593
A:Accession: S41646
A>Status: preliminary
A:Molecule type: mRNA
A:Residues: 11278 <BRO>
A:Cross-references: EMBL:L15079; NID:g310192; PIDN:AAA02937.1; PID:g310193
R:Deuchars, K.L.; Duthe, M.; Ling, Y.
Biochim. Biophys. Acta 1130, 157-165, 1992
A:Title: Identification of distinct P-glycoprotein gene sequences in rat.
A:Reference number: S22351; MUID:92223089; PMID:1348630
A:Accession: S22354
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1211-1278 <DEU>
A:Cross-references: EMBL:X61105; NID:g56892; PIDN:CAA43417.1; PID:g1334220
C:Keywords: ATP; multidrug resistance protein; ATP-binding cassette homology
C:Keywords: ATP; glycoprotein; nucleotide binding; P-loop
F:409-603/domain: ATP-binding cassette homology <ABC1>
F:426-433/region: nucleotide-binding motif A (P-loop)
F:1051-1247/domain: ATP-binding cassette homology <ABC2>
F:1068-1075/region: nucleotide-binding motif A (P-loop)

```

Query Match 74.2%; Score 4803; DB 2; Length 1278;

Best Local Similarity 73.8%; Pred. No. 6.3e-271;

Matches 948; Conservative 146; Mismatches 177; Indels 14; Gaps 8;

```

QY 1 MDPEGGRGSAEK--NFKMGKRSKKN-EKKEKKPYSTAMPFYSWMLRLVLTGMA 57
Db 1 MDLEAARNGTARLDGDEFGISISQSRKKRKNLLIGPLTFRYSMDOKRLMLGATAM 60

```

```

QY 58 AIHGAALPLMLVFEGNMTDSFA-NAGISRNKTFPVIIINESTNTNTOHFINHLEEMTY 116
Db 61 AIHAGSLPLMLVFEGNMTDFVNDAG---NFSLVPNFSLSMLNBR-----ILBEMTRY 113
QY 117 AYYSGIGGVLAAYVAYOVSFWCLAGRQILKIRKOFFHAIROEIGFVDHVDGELNTR 176
Db 114 AYYSGIGGVLAAYVAYOVSFWTLAAGRQIRKIRKOFFHAIROEIGFVDHVDGELNTR 173
QY 177 LTTDVSKEINIGDKIGMFPHSTAFFGTFYGFTRGKMLTIVILATIPVGLSAAIAK 236
Db 174 LTTDVSKEINIGDKIGMFPHSTAFFGTFYGFTRGKMLTIVILATIPVGLSAAIAK 233
QY 237 ILSFTPEKELLAARKGAVAEVLAIRTYIAFGQKKELEBYNRLBEAKNGIGIKAIT 296
Db 234 ILSFTPEKELLAARKGAVAEVLAIRTYIAFGQKKELEBYNRLBEAKNGIGIKAIT 293
QY 297 ANISIGAFLIIVASYALAFWYGTSLVLSSEKSTGOVLVFFSVLIGASIQOASPSTEA 356
Db 294 ANISIGAFLIIVASYALAFWYGTSLVLSSEKSTGOVLVFFSVLIGASIQOASPSTEA 353
QY 357 PANARGAAYELFKITDKPSTDSYKSGHKPDNIKGNLEFKNVHFSYPSRKREYKILGLN 416
Db 354 PANARGAAYELFKITDKPSTDSYKSGHKPDNIKGNLEFKNVHFSYPSRKREYKILGLN 413
QY 417 LKVGSGTVALVNGSGGKSTTVOLMQRLYDPDGAVCIDGODIRPTLNRLHREITGVYS 476
Db 414 LKVGSGTVALVNGSGGKSTTVOLMQRLYDPDGAVCIDGODIRPTLNRLHREITGVYS 473
QY 477 OEPVLPATTIENIRYGENYNTMEIEKAVEANAAYFIKMLPNKPEDTLGERSAQLSG 536
Db 474 OEPVLPATTIENIRYGENYNTMEIEKAVEANAAYFIKMLPNKPEDTLGERSAQLSG 533
QY 537 OKORITAIARALVRNPKIILLDEATSLDTESEAVVQVALDKARKGRTTYIAHRLSTVRN 596
Db 534 OKORITAIARALVRNPKIILLDEATSLDTESEAVVQVALDKARKGRTTYIAHRLSTVRN 593
QY 597 ADVIAGFDGCVYVEKGNDELAKKEGYEFLVTMOTRGNTELEENATGESKSESDELSMS 656
Db 594 ADVIAGFDGCVYVEKGNDELAKKEGYEFLVTMOTRGNTELEENATGESKSESDELSMS 653
QY 657 PKDSSSLIKRSTRSIRIAPQGDRLGTR-EDLENVNPVPSFWRIKLKNTMPTFVY 715
Db 654 P-NGMKARLFRNSTKSLKSSRAHNRDVENNELDANVPVPSFLKVLRLKRTMPTFVY 712
QY 716 GIPCALINGLOPASSTIFSRILGIFTRDEDEPTKRONSNFVLYFLVGIISFTFELQ 775
Db 713 GIPCALINGLOPASSTIFSRILGIFTRDEDEPTKRONSNFVLYFLVGIISFTFELQ 771
QY 776 GTFEFGKAGELLITRLRYVFRSMLRODYSWFPDPRNTGALTTRLANDAAGKATGSR 835
Db 772 GTFEFGKAGELLITRLSRMAFKAMLRQDMSPDDYKNSGALSTRLATDRAQVOGATG 831
QY 836 AVITQNTANTGCTGIISLIYGMQTLTLLLAIVPIIAIAGVEMKMLSGOALKDKRELEGA 895
Db 832 AVITQNTANTGCTGIISLIYGMQTLTLLLSVPFIASGIVEMKMLAGNAKDKRELEGA 891
QY 896 GKATEAIENFRVYSLTRROKFEYMAOSLOVPYNSLRKAHIFGVSTIQAMMYFSY 955
Db 892 GKATEAIENFRVYSLTRROKFEYMAOSLOVPYNSLRKAHIFGVSTIQAMMYFSY 951
QY 956 AGCRRFAGYIVANFNFDVLVFSALVIGAMAVGVSSFPADYAKAVSAAHVIMITE 1015
Db 952 AGCRRFAGYIVANFNFDVLVFSALVIGAMAVGVSSFPADYAKAVSAAHVIMITE 1011
QY 1016 KSPILDSYSPHGLKPNLTGCVTFNEVYVFPTRPDIPYLOGISLEVKKGOTIALVGSSG 1075
Db 1012 KSPILDSYSPHGLKPNLTGCVTFNEVYVFPTRPANPVLQGLSLEVKKGOTIALVGSSG 1071
QY 1076 CGKSTVVOGLLREYDPLAGSVLIDGKEIKHLNVQMLRAHLGIYSOEPILEDCSIAENIAY 1135
Db 1072 CGKSTVVOGLLREYDPLAGSVLIDGKEIKHLNVQMLRAHLGIYSOEPILEDCSIAENIAY 1131

```


C: Superfamily: multidrug resistance protein; ATP-binding cassette homology
C: Keywords: glycoprotein; membrane protein; nucleotide binding; P-loop
E: 455-462/Region: nucleotide-binding motif A (P-loop)

Query Match	49.68;	Score 3209.5;	DB 2;	Length 1321;

QY 20 KSKRNEK-KEKKPVSTPFAMFRYSNMULDRILMYLGVMAALHGAALPLMLVGNMSTDS 78
 30 KKSRLÖDKKEGDIRVGFELFRFSSKRDIMLMLGGCCLHBMADPGILLIFIGMTDI 89
 QY 79 FNAAGISR-----NKTFPVILNES---ITNNTOHFINHLEEMTTAAVYVSGI 123
 Db 90 FIITYIEROELEIPGACYNNTI-VINSSFHONNTNCTVCGLVADISEMIKFGSIGIAY 148
 Db 124 GAGVLVAAAYIOVSFWCLAAGRQILKIRKOFHAIKROETIGMFVDYHVDGELNTRLTDVSK 183
 149 GMVLLIGYQIURLWYITGARQIRMRKTYERRIMKEIGMFDCTSVGLNLSRADIIEK 208
 QY 184 INEGIDGICGMFHSIATFETFGIVGSGMKLTLYIIAISPLGSLSPAIKAKLISFSD 243
 Db 209 INAIADQULAHFÖMRSTAMWCGILLISFTYHGMKLTLYIIAIVSPILGICGAIVGLSIANFTE 268
 QY 244 KELLAYKAGAVAEVYLAARIVYIAFGGOKKELERYNKNLEBAKGIGIKATIANISIGA 303
 Db 269 LELKAYAKKASINDIVLSSIRTYAANGKNEVEYERKMLVPAQRMWGMKGVMVGFYGY 328
 QY 304 AFLIIIVASYALAEWYGTSLV--SSEISIQULVTFPSVYLIGAFSTIGQASPISEAFANARG 362
 Db 329 MGCILIFCYALAEWYGTSLVLEDEEYTPGLVQIFLCVLIAMNIGHASSCLIEFSTGCS 388
 QY 363 AAYELIEKILDNKPSIDYSKSGKPNIKGNLEFKWVHSSYPSRKEVYLKGLNLKVQSG 422
 Db 389 AATNIFÖTIDRÖPVIDCMGSDGYKIDIRKIGEIEFNHVTYHYSRBDVKIIDLMLSVYIKRG 448
 QY 423 QTVALVNGSGCKSTVOLMÖRLYDPTDGCWICIDGODITIVNRIHRELTGYVSOEPLVF 482
 Db 449 EFTVLVNGSSGAKSTALOTIORFEDPCGKWVTLGDHDIRSLNIRWLRIIDGIGVEDEPVLF 508
 QY 483 ATTIAENIRYGRENVWMDIEKAVKBAANVADFTMKLPNKEDTLYGEBRGALSGGOKORIA 542
 Db 509 STTIAENIRFREDAVTMEDIVÖAKKDINAFTALPPOORDTLVGGGGOMGGOKÖRYA 568
 Db 543 IARALVRNPKILLIDEMATSAIDTSESAVVOVALDKARKGTTIVYAHRLSTYRANADYVAG 602
 569 IARALIRNPKILLIDMATSAIDNESAARVOEALINKIOHGHTIISVAHRLSTYRAADYVIG 628
 QY 603 FDDGYIVERNHDELMEKKGITYFKLYVMQIRGNTELENA-TQESSESADLEEM-----S 656
 Db 629 FEHGAIVAEVGTHEELIERKGYFMYLVÖSGDNVAHETIMKIDATIEGGLTERTSRKS 688
 QY 657 PDDGSSSLIKRSTPR--SIHAP--QGDGDKLSTKED-----LNNVNPVPSFWRIILKLN 706
 Db 689 YRDSLRASIRÖKSÖKSÖLSLILHDPPLVAADHKSSYKOSKONDVLYEVEBPAPYRILIKYN 748
 QY 707 STEMPFVYVINGICAILNGLOPAPSILFSRIIGIFRDEDEBETIKRÖNSNMFVLVYLGI 766
 Db 749 IPEMHVILVGSLSAINGAVPIYSLFSQILGTFPSL-LDKÖORSEIHSMCLFEVILGC 807
 QY 767 ISFTIFFLGGTFPGKAGELITFRILRYMVERBMLBODYSWFPDPRKNTGALTIRLANDAQ 826
 Db 808 VSIFTOFLOGYTFKASGELLTRKLPFGFKKALGODIGWFPDDLNNPVCVLTIRLANDAQ 867
 QY 827 VKAIGSRPLAVITRONIANLGTGIIISLIYGMQTLTLLALVPIIAIAGVEMKMLSQAL 886
 Db 868 VQARISQSGVMYMNSTNIIAALLIAFPFSWKLISITIFPPALISGAAYOTKMLTGFPAS 927
 QY 887 DKRKELEGAGKIATEAIEENFRVYSLTRÖKQFERYVYASLOVPRNLSRKNIHGVGESFI 946
 Db 928 ÖDRKÖALEKQOITSEALSINRTVAGIGVEGRIFAKAEVELOJSTKYIAKRANITIGLOFAP 987
 QY 947 TOAMMTFSYAGCGRFPAIVYANDEPNMFQOVLLVFSIVFGAMAAVOVSSFAADYAKAVS 1006

Db	988	SOGIAFLANAAARYGUYLAIEGLGE SHVERRYVSSVLSAATVAGTFEFYTPSYAKAKTS	1047
Qy	1007	AAHYIMIEKSPIDISPHGLKPNLTGNTVENEYVNPTRPDIPVLOGLSLEPKKQ	1066
Db	1048	AARFOLLDKRPPINYSSEAGKWNFOGKIDETDCKFTYPSRPDIQVANGLSVSNPQ	1107
Qy	1067	TALVSSGGCKSTVVOULLERFYDPLAGSVLIDGKEIKHINOMLRAHLGIVSQEPILED	1126
Db	1108	TLAVSGSGGCKSTSIDLLEFRFYDPOGVIMJGHSBKKNVIOPLANSIGIVSQEPILED	1167
Qy	1127	CSIAENIAYGCDNSRVSHHEIMQAKKANIHFTETLPEKYNTRVGDKGTOLSGGQKORI	1186
Db	1168	CSIMDNKIYGNOMKEISVERAIAAKOALHDFEMFMSPEKRYETNVGISOGLSBERGORI	1227
Qy	1187	AIAIALVROPHILLDEFTSALDTEFSKYVOEALDKRREGTCIVIAHRSTIONADLIY	1246
Db	1228	AIAIAIVNDPKILLDEFTSALDTESEKTYOTALDKAREGTCIVIAHRLSTIONSDIIA	1287
Qy	1247	VEONKVEREGTHOOLLQAQGIYFSMV	1273
Db	1288	VVSQGVVIEKGTHEKIMQKGAYYKVL	1314

RESULT 13
T42842
bile salt transport protein, ATP-dependent - rat
N:Alternate names: bile salt export pump, sister of P-glycoprotein
C:Species: Rattus norvegicus (Norway rat)
C:Date: 11-Jan-2000 #sequence_revision 11-Jan-2000 #text_change 19-Jul-2002
C:Accession: T42842
R:Gerloff, T.; Stieger, B.; Hagenbuch, B.; Madon, J.; Landmann, L.; Roth, J.; Hofmann
J. Biol. Chem. 273, 10046-10050, 1998
A:Title: The sister of P-glycoprotein represents the canalicular bile salt export pump
A:Reference number: 222272; MUID:98212048; PMID:9545351
A:Accession: T42842
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: mRNA
A:Residues: 1-1321 <GB>
A:Cross-references: EMBL:U69487; NID:q3075421; PID:q3075422; PIDN:AAC40084.1
A:Experimental source: strain Sprague-Dawley; liver
C:Genetics:
A:Gene: spgp
C:Superfamily: multidrug resistance protein; ATP-binding cassette homology
C:Keywords: ATP; duplication; glycoprotein; membrane protein; P-loop

[illegible]

```

QY 363 AAYEIKIDNKPSIDYSKSGHKPNIKNGLEKFNHVSYPKREVKILKLANIKVSG 422
  |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 389 AATNIFQIDROPVIDOMSGDQKIDRINGERIEFHVHTHYSRDVKLDMLSWVIRPG 448
  |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 423 QYVALVNSCGKSTVVOIMQRLYDPTDGMVCIIDGDIRTTINVRHLREITGVASQEPYLF 482
  |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 449 ETTALVSSGAGKSTALQIDRFREYDCEGMVLLDGDHDIRSLNI RMLRQIGIVBEPYLF 508
  |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 483 ATTIAENIRGRNVTMDIELEKAVKANAYDFIMKLPNKFPDLVGEKRAQISGGCKORIA 542
  |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 509 STIAENIRGRNVTMDIELEKAVKANAYDFIMKLPNKFPDLVGEKRAQISGGCKORIA 568
  |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 543 IARALVNRKILLDDEATSDALDTESEAVVQVADLKARKGRTTIVIAHRLSTVRNADYAG 602
  |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 569 IARALVNRKILLDDEATSDALDTESEAVVQVADLKARKGRTTIVIAHRLSTVRNADYAG 628
  |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 603 FDDGVYERGNDELKMEKGIYFKLVMTQTRNELELENA-TGESKESDALEM-----S 656
  |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 629 FEHGVAVERTHELELKERGVFEMLVTLQSGDNNAHKETSIMGKDATEGGTLETFSGS 688
  |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 657 PKDSGSSILIKRRSTR---SIHAP--QGDRKLGKED-----LNENVPYSFWRILIKLN 706
  |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 689 YRDSLEASIRORSKSQSLTLDHPPLAVADHKSSKDSKDNVDLYEEVEPAPVRRLIKYN 748
  |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 707 STEWPFYVYICAIINGSLQAPSIITSRITGIFTEDEDETRKONSNFVLYLVGI 766
  |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 749 IPEMHTILVGSLSAAINAVPTIYSLFSLQGLTFPSL-LDEEQORSEIHSWCFEVLICG 807
  |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 767 ISFTIFFLOGFTFGAGELITRRLRYMVRFSMLRODVSWPDPKNTGALTETRLANDAQ 826
  |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 808 VSIFTOFLOGYTFASGELITRRLKRFKFMKLGODIGWDDLRNNPGLVTLTRLATDAQ 867
  |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 827 VKGALGSLAVTQNIANLGTGIIISLYGQTLTLLAIVPLAIVAGVEMKLSGOAL 886
  |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 868 VQAGAGSOYGMVNSFTMIALLIAFEFSWKLSTLITTFEPFLASGAVQTMKLGFS 927
  |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 887 KDKKLEAGKATPAIEINFRVYSLTRREKQEFEMVQASIQVPRNSLRAHIFGVSPST 946
  |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 928 QDKOLEKAGQITSEALSNTITVAGIGVEGRFKAIFEVLQTSYKAVRKANITYGLCFAR 987
  |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 947 TOAMMYFSYACCFRFGAYLVANEFMNFQDVLVFSALVFAMAVGOVSPAPPAKAKS 1006
  |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 988 SOGIAFLANSAAHYRGYLAIEGLGSHVFRVSSVALSATAVGRTFTTSPYAAKAKS 1047
  |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 1007 AAHYMIIEKSPILDSYSPHGLKNTLEGVNTNEVFNTPRPDIPLVQLSLEYKKGO 1066
  |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 1048 AARFFQLDRKPPITNVYSEAGEKMDNFQKIDFIDCKFTYPSRPDIQVNLGSVNPQ 1107
  |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 1067 TLALVSSGGCKSTVYOLLREFYDPLAGSVLIDGKEIKHLNVQMLAHGIVSQEPILFD 1126
  |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 1108 TLALVSSGGCKSTVYOLLREFYDPLAGSVLIDGKEIKHLNVQMLAHGIVSQEPILFD 1167
  |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 1127 CSTAENIYAGNSRVSHHEBIMQAEANIHPIETLPEKYNTRVGDQTSQSGCKORI 1186
  |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 1168 CSTMDNTRYDNTKEISVEFAIAAKOQLHDFVMSLPKEYENNGIQSSQ-SREKQRT 1227
  |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 1187 AIAALVROPHILLDDEATSDALDTESEAVVQVADLKARKGRTTIVIAHRLSTVQADLV 1246
  |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 1228 AIAALVROPHILLDDEATSDALDTESEAVVQVADLKARKGRTTIVIAHRLSTVQADLV 1287
  |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 1247 VFONGKVEHGHQOOLAKGIYFSMV 1273
  |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 1288 VVSQGVVIEKGTHEKIMAKGAYKLV 1314
  |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:

```

RESULT 14

hypothetical protein C47A10.1 - Caenorhabditis elegans
 C:Species: Caenorhabditis elegans
 C:Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 31-Jan-2000
 C:Accession: T19982, T24202
 R: Basham, V.

submitted to the EMBL Data Library, November 1996
 A:Reference number: Z19207
 A:Accession: T19982
 A:Status: preliminary; translated from GB/EMBL/DBJ
 A:Molecule type: DNA
 A:Residues: 1-1294 <N1294>
 A:Cross-references: EMBL:281484; PIDN: CAB03973.1; GSPDB: GN00023; CESP: C47A10.1
 A:Experimental source: clone C47A10
 R: Basham, V.
 submitted to the EMBL Data Library, April 1997
 A:Reference number: Z19853
 A:Accession: T24202
 A:Status: preliminary; translated from GB/EMBL/DBJ
 A:Molecule type: DNA
 A:Residues: 1-1294 <N1294>
 A:Cross-references: EMBL:293782; PIDN: CAB07855.1; GSPDB: GN00023; CESP: C47A10.1
 A:Experimental source: clone R12G8
 C: Genetics:
 A:Gene: CESP: C47A10.1
 A:Map position: 5
 A:Introns: 15/2; 67/3; 123/3; 168/2; 509/1; 649/3; 826/2; 1089/3; 1158/3; 1250/1
 C: Superfamily: multidrug resistance protein; ATP-binding cassette homology

Query Match 44.1%; Score 2857; DB: 2; Length 1294;
 Best Local Similarity 45.6%; Pred. No. 8e-158;
 Matches 579; Conservative 239; Mismatches 422; Indels 30; Gaps 9;

```

QY 22 SKNKKRKKPTVSTFAMFRYSNMILRLVGMVMAIINGALPLMLVFGNMVTSFAN 81
  |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 17 SKKKEAPPPKTSIFQIYRTSVDRMLAVGIVSCANGVGLPLMSIMGVNSQNFY 76
  |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 82 AGISRNKTFPVYINVESTINNTQ-----HFINHLEEMFTTAYIYSGIGAVLYAAIYVS 136
  |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 77 LG-----TFIDPNSTASKEAARAARFESHEVYONCLAKVY---IGCGIFAAFGLOAS 125
  |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 137 FWCLAAGRIILKIRKOFPHAIIMROELGPFVDHVDGELNTRLDGVSCKINGIDKGMF 196
  |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 126 CFMVICEKLSNRFRRPFHSHVMBQEIAMYDKNKGSLTSMKLPNLERVGRGTDKGLAF 185
  |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 197 HSIATFFGIVGTTRGKLTLYLAISPVGLSAAITMAKIISSFTDKELAVAKAAYA 256
  |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 186 QMNAQFVGGEFAVATYVYMLTLTIMMSLSPMMICGLFLAKTLATTAATKEKQAVAVGIA 245
  |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 257 EEVLAIETVAFGQKKELEERYKNLEAKGIGIKKAITANISGAALPLIYASTALAF 316
  |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 246 EEVLSTIRTVIAFNGQYIECKRDEALEHCKTKGKISFLIGAGLASFEVVIYASCLAF 305
  |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 317 WYGSLSVSEYSIGVLTVEFVSVLIGAFSIGQASPSIAEFANARGAAEYFIKIIDNKS 376
  |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 306 WVGTFNPFYSGRLSEGTVLVTFEFSVMGSMALGOAGQGFATIGALGAASLYEVDRIPE 365
  |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 377 IDYSKSGIHKPDNKGLEKFNHVSYPKREVKILKLANIKVSGQYVALVNSGCGKS 436
  |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 366 IDASTGQPSKISGIVSNKVEFTYPRADVKILKGVSDLAOPQYVALVNSGCGKS 425
  |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 437 TYVOLMQRLYDPTDGMVCIIDGDIRTTINVRHLREITGVASQEPYLFATIAENIRYGEN 496
  |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 426 TYVOLMQRLYDPTDGMVCIIDGDIRTTINVRHLREITGVASQEPYLFATIAENIRYGEN 485
  |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 497 VTMDIELEKAVKANAYDFIMKLPNKFPDLVGEKRAQISGGCKORIAIARALVNRKILL 556
  |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 486 VSDIEDIARALKEMADFTKTPEEGINTLVGRGYOMSGCKORIAIARALVNRKILL 545
  |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 557 DEATSDALTESEAVVQVADLKARKGRTTIVIAHRLSTVRNADYAGFDDGVYERGNHDE 616
  |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 546 DEATSDALTESEAVVQVADLKARKGRTTIVIAHRLSTVRNADYAGFDDGVYERGNHDE 605
  |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 617 LMEKGIYFKLVMTQTRNELEENATGESKESDALKMSPKDSGSSILIKRRSTR----- 671
  |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 606 LIEQKILYHELVAQVFA---DVDDKPKKEAEREMSRQTSQBRGVSVFKIQEQQVDEPK 662
  |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 672 RSIHAPQGDRLKG--TKEDLNENVPYSFWRILKLNSTEMPYVVGIFCAIINGSLQPA 729
  |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:

```



```
Db 1118 SLRQVCTVSOEPTTLDCTIGENICGTN-RNWTQELVEAAMNHNFIIGLDPGYDT 1176
Oy 1170 RVGDKGTQLSGGQKRIARALVROPHILLDEATSALDPSEKVVQEALDKAREGRTC 1229
Db 1177 HVGEKGTQLSGGQKRIARALVRSVLLIDENATSALDPSEKIVQEALDAKOGRTC 1236
Oy 1230 IVIAHRLSTIONADLIVFQNGKVEKEHGTTHOOLLAOKGIYFSMVSVQ 1276
Db 1237 LVIAHRLSTIONSDVIAIYSEKIVKEGTHDELIRKSEIYOKFCETQ 1283
```

Search completed: December 9, 2002, 17:09:31
Job time : 646 secs

GenCore version 5.1.3
Copyright (c) 1993 - 2002 Compugen Ltd.

OM protein - protein search, using sw model

Run on: December 9, 2002, 16:20:54 : Search time 290 Seconds

(without alignments)
183.211 Million cell updates/sec

Title: US-09-672-725C-2

Perfect score: 6477

Sequence: 1 MDPGGRKGSAEKNFMKMK.....LAAQGIYSMSVQAGAKR 1281

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 112892 seqs, 41476328 residues

Total number of hits satisfying chosen parameters: 112892

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SwissProt_40.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	5861	90.5	1280	1	MDR1_HUMAN
2	5662.5	87.4	1276	1	MDR1_CRIGR
3	5633	87.0	1276	1	MDR3_MOUSE
4	5296	81.8	1276	1	MDR1_MOUSE
5	5277	81.5	1276	1	MDR2_CRIGR
6	5226.5	80.7	1277	1	MDR1_RAT
7	4905.5	75.7	1279	1	MDR3_HUMAN
8	4856	75.0	1276	1	MDR2_MOUSE
9	4818.5	74.4	1281	1	MDR3_MOUSE
10	4803	74.2	1278	1	MDR2_CRIGR
11	3300	50.9	1321	1	MDR2_RAT
12	3224.5	49.8	1321	1	AB11_HUMAN
13	3207.5	49.5	1321	1	AB11_RABIT
14	3205.5	49.5	1321	1	AB11_MOUSE
15	2739	42.3	1321	1	AB11_RAT
16	2564.5	39.6	1302	1	MDR4_CAEL
17	2494.5	38.5	1302	1	MDR3_MOUSE
18	2345	36.2	1302	1	MDR3_MOUSE
19	2277	35.2	1324	1	MDR3_MOUSE
20	2072	32.0	1280	1	MDR1_MOUSE
21	1543	23.8	1419	1	MDR1_MOUSE
22	1216.5	18.8	1290	1	MDR1_MOUSE
23	1132	17.5	1336	1	MDR1_MOUSE
24	1018.5	15.7	1323	1	MDR1_MOUSE
25	893	13.8	715	1	MDR1_MOUSE
26	892	13.8	735	1	MDR1_MOUSE
27	886.5	13.7	738	1	MDR1_MOUSE
28	835	12.9	762	1	MDR1_MOUSE
29	828	12.8	762	1	MDR1_MOUSE
30	825	12.7	766	1	MDR1_MOUSE
31	820	12.6	766	1	MDR1_MOUSE
32	815	12.6	766	1	MDR1_MOUSE
33	770	11.9	820	1	MDL2_YEAST

ALIGNMENTS

RESULT 1	ID	MDR1_HUMAN	STANDARD:	PRT: 1280 AA.
AC	P08183	MDR1_HUMAN	012755: Q14812;	
DT	01-AUG-1988	(Rel. 08, Created)		
DT	01-NOV-1997	(Rel. 35, Last sequence update)		
DT	15-JUN-2002	(Rel. 41, Last annotation update)		
DE	Multidrug resistance protein 1 (P-glycoprotein 1) (CD243 antigen).			
GN	ABCB1 OR P-glyc OR MDR1.			
OS	Homo sapiens (Human).			
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
OC	Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.			
OX	NCBI-TaxID=9606;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RA	MEDLINE-87028230; PubMed-2876781;			
RA	Chen C.-J., Chin J.E., Ueda K., Clark D.P., Pastan I., Gottesman M.M.,			
RA	Roninson I.B.;			
RT	"Internal duplication and homology with bacterial transport proteins			
RT	in the mdr1 (P-glycoprotein) gene from multidrug-resistant human			
RT	cells.";			
RL	Cell 47:381-389(1986).			
RN	[2]			
RP	SEQUENCE FROM N.A.			
RA	MEDLINE-90094448; PubMed-1967175;			
RA	Chen C.-J., Clark D.P., Ueda K., Pastan I., Gottesman M.M.,			
RA	Roninson I.B.;			
RT	"Genomic organization of the human multidrug resistance (MDR1) gene			
RT	and origin of P-glycoproteins.";			
RL	J. Biol. Chem. 265:506-514(1990).			
RN	[3]			
RP	SEQUENCE FROM N.A.			
RA	MEDLINE-97190336; PubMed-9038218;			
RA	Chen G., Duran G.E., Steeger K.A., Lacayo N.J., Jaffrezou J.P.,			
RA	Dumontet C., Sikic B.I.;			
RT	"Multidrug-resistant human sarcoma cells with a mutant P-glycoprotein,			
RT	altered phenotype, and resistance to cyclosporins.";			
RL	J. Biol. Chem. 272:5974-5982(1997).			
RN	[4]			
RP	SEQUENCE OF 1-234 FROM N.A.			
RA	Smith A., Beck C., Gibson A.;			
RA	Submitted (Aug-1997) to the EMBL/GenBank/DBJ databases.			
RN	[5]			
RP	SEQUENCE OF 178-215 AND 800-856 FROM N.A.			
RA	MEDLINE-90290529; PubMed-1972623;			
RA	Gekele V., Weger S., Probst H.;			
RT	"mdr1/P-glycoprotein gene segments analyzed from various human			
RT	leukemic cell lines exhibiting different multidrug resistance			
RT	profiles.";			
RL	Biochem. Biophys. Res. Commun. 169:796-802(1990).			
RN	[6]			
RP	SEQUENCE OF 1-23 FROM N.A.			
RA	Koka N., Tsubota J., Kakehi Y., Komano T., Gottesman M.M.,			
RA	Pastan I., Ueda K.;			
RL	Submitted (Jul-1991) to the EMBL/GenBank/DBJ databases.			
RN	[7]			

34	763	11.8	598	1	Y288_THEMA
35	760	11.7	695	1	MDL1_YEAST
36	759	11.7	1437	1	MDP5_HUMAN
37	753.5	11.6	1436	1	MDP5_MOUSE
38	745	11.5	1545	1	MDP2_HUMAN
39	743.5	11.5	1436	1	MDP5_MOUSE
40	739.5	11.4	726	1	YFX9_SCHPO
41	734	11.3	1564	1	MDP2_RABIT
42	733	11.3	1325	1	MDP4_HUMAN
43	722	11.1	685	1	MDL1_CANAL
44	718.5	11.1	703	1	TAP2_RAT
45	708	10.9	584	1	MDR4_LUACIA

O9y64 thermotoga
P33310 saccharomyc
O15440 homo sapien
O9y60 ratius norv
O92887 homo sapien
O91K5 mus musculu
O9y767 schizosacch
O28689 oryctolagus
O15439 homo sapien
P97998 candida alb
P36372 ratius norv
O9chl8 lactococcus

RP VARIANTS SER-893 AND THR-893.
 RX MEDLINE-21686803. PubMed-11829140;
 RA Saito S., Iida A., Sekine A., Mura Y., Ogawa C., Kawachi S.,
 Higuchi S., Nakamura Y.;
 RT "Three hundred twenty-six genetic variations in genes encoding nine
 members of ATP-binding cassette, subfamily B (ABCB/MDR/TAP), in the
 Japanese population.";
 RT Hum. Genet. 47:38-50(2002).
 RL
 CC -1- FUNCTION: ENERGY-DEPENDENT EFFLUX PUMP RESPONSIBLE FOR DECREASED
 CC DRUG ACCUMULATION IN MULTIDRUG-RESISTANT CELLS.
 CC -1- SUBCELLULAR LOCATION: Integral membrane protein.
 CC -1- TISSUE SPECIFICITY: EXPRESSED IN LIVER, KIDNEY, SMALL INTESTINE
 CC AND BRAIN.
 CC -1- SIMILARITY: BELONGS TO THE ABC TRANSPORTER FAMILY. MDR SUBFAMILY.
 CC -1- DATABASE: NAME-Atlas Genet. Cytogenet. Oncol. Haematol.;
 CC WWW="http://www.infobiogen.fr/services/chromancer/genet/pgy11d105.html".
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
 CC or send an email to license@isb-sib.ch).
 CC
 DR EMBL, M14758; AA059575.1; -;
 DR EMBL, M29447; AA059576.1; -;
 DR EMBL, M29444; AA059576.1; JOINED.
 DR EMBL, M29425; AA059576.1; JOINED.
 DR EMBL, M29426; AA059576.1; JOINED.
 DR EMBL, M29427; AA059576.1; JOINED.
 DR EMBL, M29428; AA059576.1; JOINED.
 DR EMBL, M29429; AA059576.1; JOINED.
 DR EMBL, M29430; AA059576.1; JOINED.
 DR EMBL, M29431; AA059576.1; JOINED.
 DR EMBL, M29432; AA059576.1; JOINED.
 DR EMBL, M29433; AA059576.1; JOINED.
 DR EMBL, M29434; AA059576.1; JOINED.
 DR EMBL, M29435; AA059576.1; JOINED.
 DR EMBL, M29436; AA059576.1; JOINED.
 DR EMBL, M29437; AA059576.1; JOINED.
 DR EMBL, M29438; AA059576.1; JOINED.
 DR EMBL, M29439; AA059576.1; JOINED.
 DR EMBL, M29440; AA059576.1; JOINED.
 DR EMBL, M29441; AA059576.1; JOINED.
 DR EMBL, M29442; AA059576.1; JOINED.
 DR EMBL, M29443; AA059576.1; JOINED.
 DR EMBL, M29444; AA059576.1; JOINED.
 DR EMBL, M29445; AA059576.1; JOINED.
 DR EMBL, M29446; AA059576.1; JOINED.
 DR EMBL, AF016535; AB069423.1; -;
 DR EMBL, AC002457; AAC82531.1; -;
 DR EMBL, M37724; AAA88047.1; -;
 DR EMBL, M37725; AAA88047.1; -;
 DR EMBL, X58723; CAA15558.1; -;
 DR PIR: A25059; DVHUI.
 DR PIR: A34914; A34914.
 DR Genev; HGNC:40; ABCB1.
 DR MIM: 171050; -;
 DR InterPro: IPR003593; AAA_ATPase.
 DR InterPro: IPR003439; ABC_transporter.
 DR InterPro: IPR001140; ABCtransprtTM.
 DR Pfam: PF00005; ABC_tran. 2.
 DR Pfam: PF00664; ABC_membrane. 2.
 DR ProDom: PD000006; ABC_transporter. 2.
 DR SMART: SMO0382; AAA. 2.
 DR PROSITE: PS00211; ABC_TRANSPORTER. 2.
 DR ATP-binding: Glycoprotein; Transmembrane; Transport; Repeat;
 KM Multigene family; Polymorphism.
 KW Domain 1 51 CYTOPLASMIC (POTENTIAL).
 FT TRANSMEM 52 72 POTENTIAL.
 FT TRANSMEM 120 140 POTENTIAL.
 FT TRANSMEM 189 209 POTENTIAL.

FT	TRANSMEM	216	236	POTENTIAL.
FT	TRANSMEM	297	317	POTENTIAL.
FT	TRANSMEM	326	346	POTENTIAL.
FT	DOMAIN	347	710	CYTOPLASMIC (POTENTIAL).
FT	TRANSMEM	711	731	POTENTIAL.
FT	TRANSMEM	757	777	POTENTIAL.
FT	TRANSMEM	833	853	POTENTIAL.
FT	TRANSMEM	854	874	POTENTIAL.
FT	TRANSMEM	937	957	POTENTIAL.
FT	TRANSMEM	974	994	POTENTIAL.
FT	DOMAIN	995	1280	CYTOPLASMIC (POTENTIAL).
FT	NP_BIND	427	434	ATP (BY SIMILARITY).
FT	NP_BIND	1070	1077	ATP (BY SIMILARITY).
FT	REPEAT	1	637	
FT	REPEAT	638	1280	
FT	CARBOHYD	91	91	N-LINKED (GLCNAC. . .) (POTENTIAL).
FT	CARBOHYD	94	94	N-LINKED (GLCNAC. . .) (POTENTIAL).
FT	CARBOHYD	99	99	N-LINKED (GLCNAC. . .) (POTENTIAL).
FT	VARIANT	893	893	A -> S (IN DBSNP:2032562).
FT	VARIANT	893	893	/FTID=VAR_013361.
FT	VARIANT	893	893	A -> T.
FT	CONFLICT	23	23	/FTID=VAR_013362.
FT	CONFLICT	185	185	S -> R (IN REF. 6).
FT	CONFLICT	336	336	G -> V (IN REF. 1 AND 3).
FT	CONFLICT	412	412	MISSING (IN REF. 3).
FT	CONFLICT	438	438	G -> A (IN REF. 3).
FT	CONFLICT	438	438	Q -> S (IN REF. 3).
FT	SEQUENCE	1280 AA;	141462 MW;	ABJC279531F43675 CRC64;

Query Match 90.5%; Score 5861; DB 1; Length 1280;
 Best local similarity 90.7%; Pred. No. 0;
 Matches 1163; Conservative 55; Mismatches 60; Indels 4; Gaps 4;

QY	1	MDPBGKRGSA-EKNEWKMGKSKKNEKKKPYVTFPAFRISNMIDRLMYLVTMAI	59
DB	1	MDLEDNRGKAKKNEFKLNKSKER-DKREKPPYVTFPFRISNMIDRLMYLVTMAI	59
QY	60	IHGALPLMLVFGNMTDFPANGISRNKTPVYINESTINNOHFINHLEDEMTYAY	119
DB	60	IHGALPLMLVFGNMTDFPANGISRNKTPVYINESTINNOHFINHLEDEMTYAY	117
QY	120	YSGIGAGVLAAYVLOVSWFCLAAAGROIKRKOPFAIMROEIGMPYHVGELNRLTD	179
DB	118	YSGIGAGVLAAYVLOVSWFCLAAAGROIKRKOPFAIMROEIGMDYHVGELNRLTD	177
QY	180	DVSKINSGIDKIGMEPHSTAFPTGFTVGTGRKLTVLALISPYLGISAIAIMAKIS	239
DB	178	DVSKINSGIDKIGMEPHSTAFPTGFTVGTGRKLTVLALISPYLGISAIAIMAKIS	237
QY	240	SEPTKEELAAVAKAGAAVEEYLAIRTVIAFGGOKKELERYKNLEAKGIGIKKAITANI	299
DB	238	SEPTKEELAAVAKAGAAVEEYLAIRTVIAFGGOKKELERYKNLEAKRIGIKKAITANI	297
QY	300	STGAAPFLITYASYALAFWGTSLVLSSEYSIGOVLYTFEFSVLGASISGASPTSEAFAN	359
DB	298	STGAAPFLITYASYALAFWGTSLVLSSEYSIGOVLYTFEFSVLGASISGASPTSEAFAN	357
QY	360	ARGAAYELFKTIDNKPSIDSKSGKRPDNITKNLEKRNHVFSPSKREKYLKGLNLKY	419
DB	358	ARGAAYELFKTIDNKPSIDSKSGKRPDNITKNLEKRNHVFSPSKREKYLKGLNLKY	417
QY	420	QSGOTVALVNSGSGKSTVOLMORLYPTDQVNCIDGODIRTIYVNRHLEITGVVSOEP	479
DB	418	QSGOTVALVNSGSGKSTVOLMORLYPTDQVNCIDGODIRTIYVNRHLEITGVVSOEP	477
QY	480	VLPFTTAENIRYGRNSVYIMDELEKAYVEANAAYDFIKLNKCDTLVGRGQOLSGGOKO	539
DB	478	VLPFTTAENIRYGRNSVYIMDELEKAYVEANAAYDFIKLNKCDTLVGRGQOLSGGOKO	537
QY	540	RIAIARLVNPKITLLDEATSDLTDESEAVYVALDKARKKRTTVIAHRLSTVRADY	599
DB	538	RIAIARLVNPKITLLDEATSDLTDESEAVYVALDKARKKRTTVIAHRLSTVRADY	597


```

Db 60 HGVALLPMLVFGMDTDSFASVGNIPINAT-----NNATGVNSDIEJGKLEEMETTYAYY 114
Oy 120 YSGIGAVLVAAVYIOVSFWCLAAQRQILKIRKOFFHAIKMOEIGMPYHVHVGELNRLTD 179
Db 115 YGIGAGVLIYAVYIOVSFWCLAAQRQILKIRKOFFHAIKMOEIGMPYHVHVGELNRLTD 174
Oy 180 DYSKINEGIGKIGMFEHSIATFEFGFVGTGKMLTVLILASIPVLGSAAMAKILS 239
Db 175 DYSKINEGIGKIGMFEHSIATFEFGFVGTGKMLTVLILASIPVLGSAAMAKILS 234
Oy 240 SFTDELLAYAKAGAVAEVLAIRTVIAFGOKKELEERINKLEPAKIGIKKAITANI 299
Db 235 SFTDELLAYAKAGAVAEVLAIRTVIAFGOKKELEERINKLEPAKIGIKKAITANI 294
Oy 300 SIGAFLILYASALAFWYGTSLVSSSEYSGOVLTVFESVLIGAFSIGOASPIEAFAN 359
Db 295 SMOAFLILYASALAFWYGTSLVSKESYGOVLTVFPAVLIGAFSIGOASPIEAFAN 354
Oy 360 ARGAAEIKIIDNKPISIDYSKSGHKPDNIKGNLEFNKVSFYSRKEVKILKGLMKV 419
Db 355 ARGAAEIKIIDNKPISIDYSKSGHKPDNIKGNLEFNKVSFYSRKEVKILKGLMKV 414
Oy 420 QSGQIVAVVNSGCGKSTTVOLMORLYDPTDGMVCIDODIRTNVRLREITGVSOEP 479
Db 415 QSGQIVAVVNSGCGKSTTVOLMORLYDPTDGVASIDODIRTNVRLREITGVSOEP 474
Oy 480 VLEATTIAENIRYGRENVYMDLEKAVKANAYDFIMKLPNKFDVLVGERGAOLSGGOK 539
Db 475 VLEATTIAENIRYGRENVYMDLEKAVKANAYDFIMKLPHKFDVLVGERGAOLSGGOK 534
Oy 540 RIAIARALVNPRIKILLDETSALDTESEPAVVOVALDKARKRTTIVIAHRLSTYRNDV 599
Db 535 RIAIARALVNPRIKILLDETSALDTESEPAVVOVALDKARKRTTIVIAHRLSTYRNDI 594
Oy 600 IAGFDGCVIYKGNHDELMKEKITYKLTVMOTRGNEIELENAATESSESDALEMSKD 659
Db 595 IAGFDGCVIYKGNHDELMKEKITYKLTVMOTRGNEIELENAATESSESDALEMSKD 654
Oy 660 SSSSLIKRRSTRSIAHAPOGDRKLTGKEDLNEENPVPSFWKILKINSTEWEYFVVGIFC 719
Db 655 SASSLIRRSSTRSIRSGPHDRLSTKEALDEDPVPISEFWRLKLNSEWEYFVVGIFC 714
Oy 720 AINGGIQAPASITFSRITGIFTRDDDEKRONSMFSLVFIAGISIFTFEFGFTE 779
Db 715 AIVGALQAPASITFSRIVSVFTRNDDEKRDMSLFIILIGVISITFIFFGFTE 774
Oy 780 GKAGEILTKRLRYWFRSMLRDVSFDEPKNTGALTTLRLANDAQVKAIGSLAVIT 839
Db 775 GKAGEILTKRLRYWFRSMLRDVSFDEPKNTGALTTLRLANDAQVKAIGSLAVIT 834
Oy 840 QNANLGTGIIISLYQMQLTLLAIVPILAIAGVEMKMSGOLKXKKELEGAKTA 899
Db 835 QNANLGTGIIISLYQMQLTLLAIVPILAIAGVEMKMSGOLKXKKELEGAKTA 894
Oy 900 TEATENFTVSLTRBQKFEFMYAOSLOVYRNSLRKAHIFGVFSITQAMTFYAGCF 959
Db 895 TEATENFTVSLTRBQKFEFMYAOSLOVYRNSLRKAHIFGVFSITQAMTFYAGCF 954
Oy 960 RFGATLVANEFMNFQDVLVFSATVFGAMAVGVSSFADPYAKAKYSAHVIITKSP 1019
Db 955 RFGATLVANEFMNFQDVLVFSATVFGAMAVGVSSFADPYAKAKYSAHVIITKSP 1014
Oy 1020 IDSISPHGLKPTLEGNTENEVVENVPTRPDIPLYLOGISLEVKKGOTLAVSSGCGKS 1079
Db 1015 IDSISPHGLKPTLEGNTENEVVENVPTRPDIPLYLOGISLEVKKGOTLAVSSGCGKS 1074
Oy 1080 TYVOLLEFYPDIAGSVLIDKEIKHLNVMQLRAHLGIVSOEPLDFDCAETAINAYGDN 1139
Db 1075 TYVOLLEFYPDIAGSVLIDKEIKHLNVMQLRAHLGIVSOEPLDFDCAETAINAYGDN 1134
Oy 1140 RVVSHHEIMQAKKANIHFIETLPEKNTVSGKQVLSGGOKQRIAIRALVROPHIL 1199
Db 1135 RVVSHHEIMQAKKANIHFIETLPEKNTVSGKQVLSGGOKQRIAIRALVROPHIL 1194

```

```

Db 1135 RVVSHHEIMQAKKANIHFIETLPEKNTVSGKQVLSGGOKQRIAIRALVROPHIL 1194
Oy 1200 LIDEATSAIDTESERKVOEALDKARKREGTCIVIAHRLSTIONADLIVONGKVKHGT 1259
Db 1195 LIDEATSAIDTESERKVOEALDKARKREGTCIVIAHRLSTIONADLIVONGKVKHGT 1254
Oy 1260 QOLLAOKGIYSVSVQAGAKR 1281
Db 1255 QOLLAOKGIYSVSVQAGAKR 1276

RESULT 3
MR3 MOUSE
ID MR3 MOUSE STANDARD; PRT: 1276 AA.
AC P21477.
DT 01-MAY-1991 (Rel. 18, Created)
DT 01-APR-1993 (Rel. 25, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Multidrug resistance protein 3 (P-glycoprotein 3) (MDR1A).
GN ABCB1A OR ABCB4 OR PGY3 OR PGY-3 OR MDR3 OR MDR1A.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=90287150; PubMed=1969610;
RA Hsu S.I.H., Cohen D., Kirschner L.S., Lothstein L., Hartslein M.,
RA Horwitz S.B.,
RT "Structural analysis of the mouse mdr1a (P-glycoprotein) promoter
RT reveals the basis for differential transcript heterogeneity in
RT multidrug-resistant J774.2 cells."
RL MOL. Cell. Biol. 10:3596-3606(1990).
RN [3]
RP SEQUENCE OF 173-1276 FROM N.A.
RC STRAIN=BA1B/C;
RX MEDLINE=89308614; PubMed=2473069;
RA Hsu S.I.H., Lothstein L., Horwitz S.B.,
RT "Differential overexpression of three mdr gene family members in
RT multidrug-resistant J774.2 mouse cells. Evidence that distinct P-
RT glycoprotein precursors are encoded by unique mdr genes."
RL J. Biol. Chem. 264:12053-12062(1989).
CC -1- FUNCTION: ENERGY-DEPENDENT EFFLUX PUMP RESPONSIBLE FOR DECREASED
CC DRUG ACCUMULATION IN MULTIDRUG-RESISTANT CELLS.
CC -1- SUBCELLULAR LOCATION: Integral membrane protein.
CC -1- MISCELLANEOUS: IN MOUSE THE MDR GENE FAMILY INCLUDES THREE OR MORE
CC RELATED BUT DISTINCT CELLULAR GENES.
CC -1- SIMILARITY: BELONGS TO THE ABC TRANSPORTER FAMILY. MDR SUBFAMILY.

-----
CC CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sdb.ch/announce/
CC or send an email to license@sdb-sdb.ch).
CC -----
DR EMBL: M30697; AAA39517.1; -
DR EMBL: M33581; AAA39514.1; -
DR EMBL: M33580; AAA39518.1; -
DR EMBL: M24417; AAA03243.1; -
DR PIR: A34175; DWSM1A.
DR PIR: A34786; A34786.
DR MGI: MGI:97570; Abcb1a.
DR InterPro: IPR003593; AAA_ATPase.
DR InterPro: IPR003439; ABC_transporter.

```


[illegible]

Db	474	VFATTAENIRYGRVDYTMDEIEKAVENAYDFIMKLPHQDFTLVYRGASQLSGGQK	533
QY	540	RIAIRALVNRPKILLIDDEATSAIDTESAVVVALDKARKERTIVIAHRLSTVRADY	599
Db	534	RIAIRALVNRPKILLIDDEATSAIDTESAVVVALDKARKERTIVIAHRLSTVRADY	593
QY	600	INGPDGYIVKGNHDELMKEKGYTEKLVYTMQTRGNELELENATGESSEDALEMSPKD	659
Db	594	IAGFGGVIVGQHNHDELMKEKGYTEKLVYTMQTAGNEILELGNACKSKDEIDNIDMSKD	653
QY	660	SGSSLIRKRRSRRIAHPOQDRKRLTGKEDLNNPVPVSFPRILLKLNSTEMPVYVIGIFC	719
Db	654	SGSSLIRRRSTRKSGCPHDHODRKLETKRALDEDPVPSFPRILLKLNSTEMPVYVIGIFC	713
QY	720	AIINGLOPAFTISFKILIGIFTRDEDPETKRQNSNMFSVLEVLGIIISFTFLOGTFE	779
Db	714	AIINGLOPAFTISFKIVSVFSGVVFETNGGPPETQRQNSNLSLFLILIGIISFTFLOGTFE	773
QY	780	GKAGEILLTKRLRVNVRFSMLRODVSEFDDPPKNTGATLTRLANDAAVYKATISRLAVIT	839
Db	774	GKAGEILLTKRLRVNVRFSMLRODVSEFDDPKNTGATLTRLANDAAVYKATISRLAVIT	833
QY	840	ONINLCTGIIISLITGMQTLTLLLAIVPIITAIAGVEMKMLSGQALKDKELEGAGKIA	899
Db	834	ONINLCTGIIISLITGMQTLTLLLAIVPIITAIAGVEMKMLSGQALKDKELEGAGKIA	893
QY	900	TEALENFRVYSLTRBQKFEVYMAQSLQVPRNSLRKAHIEGVSFTIQQMMYFAGCF	959
Db	894	TEALENFRVYSLTRBQKFEVYMAQSLQIPYRNKMKAHNFGITFEFTQMMYFAGCF	953
QY	960	RFAYIVLVPFNNFOVLIVYSALYFGAMAVGVSSPAPDYAKAKVSAHVIMIEKSPIL	1019
Db	954	RFAYIVLQQLMFTFENVLIVYSALYFGAMAVGVSSPAPDYAKAKVSAHVIMIEKSPIL	1013
QY	1020	IDTSYSHGLKPNLTBEGNATFENEVENFNPPTBPIPVYQGLSLKVKKQQTALVGGSGGCKS	1079
Db	1014	IDTSYSHGLKPNLTBEGNATFENEVENFNPPTBPIPVYQGLSLKVKKQQTALVGGSGGCKS	1073
QY	1080	TYVOLLERPYDPLASVYLIDKEIKELHNQWMLRAHLGIVSOEPIIFDCSIAENIAYGDN	1139
Db	1074	TYVOLLERPYDPMAGSVFLIDKEIKELQNLQWMLRAOGLVISOEPIIFDCSIAENIAYGDN	1133
QY	1140	RYVSHBEIMQAARKEANIHHTLETLPEKYNTRVGDKQTQSGGQKRIIAALVROPHIL	1199
Db	1134	RYVSHBEIMQAARKEANIHHTLETLPEKYNTRVGDKQTQSGGQKRIIAALVROPHIL	1193
QY	1200	LIDDEATSAIDTESKRYVVALDKARKERTIVIAHRLSTIONADLIYVONGKVEKHTH	1259
Db	1194	LIDDEATSAIDTESKRYVVALDKARKERTIVIAHRLSTIONADLIYVONGKVEKHTH	1253
QY	1260	QQLLAOKGITYSMVSVQAGAKR 1281	
Db	1254	QQLLAOKGITYSMVSVQAGAKR 1275	
RESULT 4			
MDRL_MOUSE	ID	MDRL_MOUSE	STANDARD; PRT: 1276 AA.
AC	P06795;		
DT	01-JAN-1988 (Rel. 06, Created)		
DT	15-JUN-1988 (Rel. 06, Last sequence update)		
DT	01-JUN-2002 (Rel. 41, Last annotation update)		
DE	Multidrug resistance protein 1 (P-glycoprotein 1).		
GN	ABCB1 OR ABCB1B OR PGY1 OR PGY1-1 OR MDRL OR MDRLB.		
OS	Mus musculus (Mouse).		
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		
OC	Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.		
OX	NCBI_TaxID=10090;		
RN	[1]		
RP	SEQUENCE FROM N.A.		
RX	MEDLINE=87028229; PubMed=3768958;		
RT	Gros P., Croop J., Housman D.;		
	"Mammalian multidrug resistance gene. Complete cDNA sequence"		

[illegible][illegible]

81.8%; Score 5296; DB 1; Length 1276;

Db 1016 IDSTYEGIKPTELEGNKFNQVNPYTRPNIPVLOGLSLEWKKGQTLALVSGSGCGKS 1075
QY 1080 TVVQLERPYDPLAGSVLIDGKEIKHLVQWLRALGLVSEPIILFDGSIENIAYGDNS 1139
Db 1076 TVVQLERPYDPLAGSVLIDGKEIKHLVQWLRALGLVSEPIILFDGSIENIAYGDNS 1135
QY 1140 RVVSHHEIYRAKEANIHFIETLEPEKYNTRYGDKGTOLSGOKORTAIRALVROPHIL 1199
Db 1136 RVVSHHEIYRAKEANIHFIETLEPEKYNTRYGDKGTOLSGOKORTAIRALVROPHIL 1195
QY 1200 ILDEATSAIDTESEKVVQEAIDKAREGTCIVIAHRLSTIQNADLIYVFGNGVKEHGT 1259
Db 1196 ILDEATSAIDTESEKVVQEAIDKAREGTCIVIAHRLSTIQNADLIYVFGNGVKEHGT 1255
QY 1260 QOLLAOKGIYFMSVQAGAKR 1281
Db 1256 QOLLAOKGIYFMS-VQAGAKR 1275

RESULT 5

MDR2_CRIGR

ID MDR2_CRIGR STANDARD; PRT; 1276 AA.

AC P2149;

DT 01-MAY-1991 (Rel. 18, Created)

DT 01-NOV-1991 (Rel. 20, Last sequence update)

DE 16-OCT-2001 (Rel. 40, Last annotation update)

GN Multidrug resistance protein 2 (P-glycoprotein 2).

OS Cricetus griseus (Chinese hamster).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Cricetinae;

OX NCBI_TaxId=10029;

[1]

RP SEQUENCE FROM N.A.

RX MEDLINE=92135896; PubMed=1685679;

RA Endicott J.A., Sarangi F., Ling V.;

RT "Complete cDNA sequences encoding the Chinese hamster P-glycoprotein

RL DNA Seq. 2:89-101(1991).

[2]

RP SEQUENCE OF 622-1276 FROM N.A.

RX MEDLINE=88122132; PubMed=2893255;

RA Endicott J.A., Utracki P.F., Sarangi F., Gerlach J.H., Deuchars K.L.,

RT Ling V.;

RT "Simultaneous expression of two P-glycoprotein genes in

RT drug-sensitive Chinese hamster ovary cells."

RL Mol. Cell. Biol. 7:4075-4081(1987).

CC -! FUNCTION: ENERGY-DEPENDENT EFFLUX PUMP RESPONSIBLE FOR DECREASED

CC -! SUBCELLULAR LOCATION: Integral membrane protein.

CC -! MISCELLANEOUS: PGP ISOFORMS DIFFER IN THEIR DRUG TRANSPORT

CC CAPABILITIES: PGP1 AND PGP2 CAN MEDIATE MDR, WHILE PGP3 APPARENTLY

CC CANNOT.

CC -! SIMILARITY: BELONGS TO THE ABC TRANSPORTER FAMILY. MDR SUBFAMILY.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration

CC between the Swiss Institute of Bioinformatics and the EMBL outstation-

CC use by non-profit institutions as long as its content is in no way

CC modified and this statement is not removed. Usage by and for commercial

CC or send an email to license@isb-sib.ch).

CC EMBL: M60041; AAA68884.1; -

DR EMBL: M1786; AAA37007.1; -

DR PIR: B27126; DVHY2C.

DR InterPro: IPR003593; AAA_Arpase.

DR InterPro: IPR003439; ABC_transportr.

DR InterPro: IPR001140; ABCtransportrTM.

DR Pfam: PF00664; ABC_tran; 2.

DR ProDom: PD000006; ABC_transportr; 2.

DR SMART: SM00382; AAA. 2.
DR PROSITE: PS00211; ABC_TRANSPORTER; 2.
KW ATP-binding; Glycoprotein; Transmembrane; Transport; Repeat;
KW Multigene family.
FT DOMAIN 1 51 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 52 72 POTENTIAL.
FT TRANSMEM 119 139 POTENTIAL.
FT TRANSMEM 188 208 POTENTIAL.
FT TRANSMEM 215 235 POTENTIAL.
FT TRANSMEM 296 316 POTENTIAL.
FT TRANSMEM 325 345 POTENTIAL.
FT DOMAIN 346 708 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 709 729 POTENTIAL.
FT TRANSMEM 755 775 POTENTIAL.
FT TRANSMEM 831 851 POTENTIAL.
FT TRANSMEM 852 872 POTENTIAL.
FT TRANSMEM 935 955 POTENTIAL.
FT TRANSMEM 972 992 POTENTIAL.
FT DOMAIN 993 1276 CYTOPLASMIC (POTENTIAL).
FT NP_BIND 426 433 ATP (POTENTIAL).
FT NP_BIND 1068 1075 ATP (POTENTIAL).
FT REPEAT 1 635
FT REPEAT 636 1276
SQ SEQUENCE 1276 AA; 141057 MW; 5096B1385628812D CRC64;

Query Match 81.5%; Score 5277; DR 1; Length 1276;
Best Local Similarity 79.9%; Pred. No. 1,2e-290;
Matches 1024; Conservative 129; Mismatches 122; Indels 6; Gaps 3;

QY 1 MDEEGKRGKGAENEFKMKKKKKKKKKPKPVSTFAFRYSNMLDRLYLVTMAIY 60
Db 1 MDEEGKRGKGAENEFKMKKKKKKKKKPKPVSTFAFRYSNMLDRLYLVTMAIY 60
QY 61 HGAALPLMLVFGNMTDFSNAGISRNKTFPIINSESTNNTQHTINLLEEMTYAYY 120
Db 61 HGAALPLMLVFGNMTDFSNAGISRNKTFPIINSESTNNTQHTINLLEEMTYAYY 120
QY 121 SGIGAGVIAAATVQVSPFCLAGROILIRKOFPAIRKOEIGFVDVDELTRTLTD 180
Db 121 SGIGAGVIAAATVQVSPFCLAGROILIRKOFPAIRKOEIGFVDVDELTRTLTD 180
QY 118 TGIAGVILVATVQVSPFCLAGROILIRKOFPAIRKOEIGFVDVDELTRTLTD 177
Db 118 TGIAGVILVATVQVSPFCLAGROILIRKOFPAIRKOEIGFVDVDELTRTLTD 177
QY 181 VSKINEGIDGKIGMFPFHSIATFETFTYFTGRKTLVILATISPVGLSAAIYAKLSS 240
Db 181 VSKINEGIDGKIGMFPFHSIATFETFTYFTGRKTLVILATISPVGLSAAIYAKLSS 240
QY 178 VSKINEGIDGKIGMFPFHSIATFETFTYFTGRKTLVILATISPVGLSAAIYAKLSS 237
Db 178 VSKINEGIDGKIGMFPFHSIATFETFTYFTGRKTLVILATISPVGLSAAIYAKLSS 237
QY 241 FTKKELLYAARAGAAEVLAIKRTVIAFGQKKLEIRYKNLEAKGIGIKKAITANIS 300
Db 241 FTKKELLYAARAGAAEVLAIKRTVIAFGQKKLEIRYKNLEAKGIGIKKAITANIS 300
QY 301 IGAFLVLYASVYALAFWYGTSLVSESTIGOVITVFVSVLIGAFSIGQASPEIAFANA 360
Db 301 IGAFLVLYASVYALAFWYGTSLVSESTIGOVITVFVSVLIGAFSIGQASPEIAFANA 360
QY 298 IGIATLVYASVYALAFWYGTSLVSESTIGOVITVFVSVLIGAFSIGQASPEIAFANA 357
Db 298 IGIATLVYASVYALAFWYGTSLVSESTIGOVITVFVSVLIGAFSIGQASPEIAFANA 357
QY 361 RGAAYEYFKIINDKPSIDSYSKSGHKPDNKGLEKRNKVFSPSKREYIKGLXKQ 420
Db 361 RGAAYEYFKIINDKPSIDSYSKSGHKPDNKGLEKRNKVFSPSKREYIKGLXKQ 420
QY 358 RGAAYEYFKIINDKPSIDSYSKSGHKPDNKGLEKRNKVFSPSKREYIKGLXKQ 417
Db 358 RGAAYEYFKIINDKPSIDSYSKSGHKPDNKGLEKRNKVFSPSKREYIKGLXKQ 417
QY 421 SGQTVAVLNGSGCKSTTVOLMRLDPTDGMVCIIGODIRITINVRHLREITGVYSEPV 480
Db 421 SGQTVAVLNGSGCKSTTVOLMRLDPTDGMVCIIGODIRITINVRHLREITGVYSEPV 480
QY 418 SGQTVAVLNGSGCKSTTVOLMRLDPTDGMVCIIGODIRITINVRHLREITGVYSEPV 477
Db 418 SGQTVAVLNGSGCKSTTVOLMRLDPTDGMVCIIGODIRITINVRHLREITGVYSEPV 477
QY 481 LPAITTAENIRYGRENVYMDIEKAVKANAYDFIKLKNKFDLTVGERGAOLSGGOK 540
Db 481 LPAITTAENIRYGRENVYMDIEKAVKANAYDFIKLKNKFDLTVGERGAOLSGGOK 540
QY 478 LPAITTAENIRYGRENVYMDIEKAVKANAYDFIKLKNKFDLTVGERGAOLSGGOK 537
Db 478 LPAITTAENIRYGRENVYMDIEKAVKANAYDFIKLKNKFDLTVGERGAOLSGGOK 537
QY 541 IAIARALVNRPKTLILDEATSAIDTESEAVVQALDKARKGTTIVIAHRLSTYRNADVI 600
Db 541 IAIARALVNRPKTLILDEATSAIDTESEAVVQALDKARKGTTIVIAHRLSTYRNADVI 600
QY 538 IAIARALVNRPKTLILDEATSAIDTESEAVVQALDKARKGTTIVIAHRLSTYRNADVI 597
Db 538 IAIARALVNRPKTLILDEATSAIDTESEAVVQALDKARKGTTIVIAHRLSTYRNADVI 597
QY 601 AGPDGCVIVEKGNHDLMEKGIYFKLVYMTQNGNIELENAVTGSKSRESALMSPRDS 660
Db 601 AGPDGCVIVEKGNHDLMEKGIYFKLVYMTQNGNIELENAVTGSKSRESALMSPRDS 660
QY 598 AGPDGCVIVEKGNHDLMEKGIYFKLVYMTQNGNIELENAVTGSKSRESALMSPRDS 657
Db 598 AGPDGCVIVEKGNHDLMEKGIYFKLVYMTQNGNIELENAVTGSKSRESALMSPRDS 657


```

OY 418 KVSQGVAVLVNSGCGKSTVQLMQLRYPDTPDGMVCIDGODIRTNVRLAREITGVSG 477
    |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 415 KKSQGVAVLVNSGCGKSTVQLMQLRYPDTPDPIEBEVSIDQDIRTNVRLAREITGVSG 474
OY 478 EPLVATIAINIKYGRNENTMDEIEAVKAAVYDFIMKLPRKFDPLVGERGAQLSGSQ 537
    |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 475 EPLVATIAINIKYGRNENTMDEIEAVKAAVYDFIMKLPRKFDPLVGERGAQLSGSQ 534
OY 538 KORIAIARALVRNPKILLDEATSAIDTESEAVVVALDKARGRITVIAHRLSTVRNA 597
    |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 535 KORIAIARALVRNPKILLDEATSAIDTESEAVVVALDKARGRITVIAHRLSTVRNA 594
OY 598 DVINGFDGYIVKGNHDELMKRGKGYFLVYMTQIRNGNEIELENAAGESSEDLKEMP 657
    |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 595 DVINGFDGYIVKGNHDELMKRGKGYFLVYMTQIRNGNEIELENAAGESSEDLKEMP 654
OY 658 KDSGSLIKRRSTRSRIRHPOGDRKLGKEDLENVPVSWIRILKLNSTEMPYVYGI 717
    |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 655 EESKSPIL-RSIRRSIRHRODERLSRKEVEDVAVSWOIKLINISEMPYLVYGV 713
OY 718 FCATINGLOPAPSIIFSRIGIFTRDEDEPTRKONSNSFVLVLGISTPFIFLOGF 777
    |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 714 LCATINGLOPAPSIIFSRIGIFTRDEDEPTRKONSNSFVLVLGISTPFIFLOGF 773
OY 778 TFGKAGEILLKRLRYWVFESMLRODVSWFDDPKNTTGALTFRLANDAAQVGAIGSRILAV 837
    |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 774 TFGKAGEILLKRLRYWVFESMLRODVSWFDDPKNTTGALTFRLANDAAQVGAIGSRILAV 833
OY 838 ITQNIANTLGTGIIIS--LTYGQVLTLALVPIIAIAGVYEMKLSQALKKKELEGA 895
    |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 834 VYQNVANLGTGIIISLVLVYGMQLLLVLIIFLVLGITIEMLKLSQALKKKELEIS 893
OY 896 GKATPAIENFTVVSILRQKFEYMAQSLQVPRNSLRKRAHIFGVSFSTQAMFYSY 955
    |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 894 GKATPAIENFTVVSILRQKFEYMAQSLQVPRNSLRKRAHIFGVSFSTQAMFYSY 953
OY 956 ACCEFGAYLVANEFMANODVLVSAIVFGAMAVGOVSSAPRYAAKYSAAHVMIIE 1015
    |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 954 AACPFRRGAYLVARELMTFENMLVFSAVFGAMAGNTSSAPRYAAKYSASHIIGIIE 1013
OY 1016 KSPILDSISPHGLKNTLEGVNTEVENYPTRPDIPIVLQSLILEVKKQOTLALVSSG 1075
    |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 1014 KSPILDSISPHGLKNTLEGVNTEVENYPTRPDIPIVLQSLILEVKKQOTLALVSSG 1073
OY 1076 CGKSTVYVLLREFYDPLAGSVLIDGKEIKHLNQMVLRAHLGIYSOEPIFDGSAIENIAY 1135
    |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 1074 CGKSTVYVLLREFYDPLAGSVLIDGKEIKHLNQMVLRAHLGIYSOEPIFDGSAIENIAY 1132
OY 1136 GDSRVVSHHEIQAARKEANTHPIETLPEKYNTRVGDKTGTOSSGOKORIAIARALVQ 1195
    |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 1133 GDSRVVSHHEIQAARKEANTHPIETLPEKYNTRVGDKTGTOSSGOKORIAIARALVQ 1192
OY 1196 PHILLDEATSAIDTESEKVOELDKARGRITVIAHRLSTIONADLLIVFONGKYKE 1255
    |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 1193 PHILLDEATSAIDTESEKVOELDKARGRITVIAHRLSTIONADLLIVFONGKYKE 1252
OY 1256 HGHQQLLAQKGYFSVAVQAGAKR 1281
    |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 1253 HGHQQLLAQKGYFSM--VQAGAKR 1276

```

RESULT 7

```

MDR3_HUMAN
ID MDR3_HUMAN STANDARD; PRT: 1279 AA.
AC P21439;
DT 01-MAY-1991 (Rel. 18, Created)
DT 01-MAY-1991 (Rel. 18, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Multidrug resistance protein 3 (P-glycoprotein 3).
GN ABCB4 OR PGY3 OR MDR3.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

```

```

OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=89138016; Pubmed=2906314;
RA van der Bliek A.M., Koelman P.M., Schneider C., Borst P.;
RT "Sequence of mdr3 cDNA encoding a human P-glycoprotein.";
RL gene 71:401-411(1988).
RN [2]
RP SEQUENCE OF 856-1279 FROM N.A.
RX MEDLINE=8811519; Pubmed=2892668;
RA van der Bliek A.M., Baas F., ten Houte de Lange T., Koelman P.M.,
RT "The human mdr3 gene encodes a novel P-glycoprotein homologue and
RT gives rise to alternatively spliced mRNAs in liver.";
RL EMBO J. 6:3325-3331(1987).
RN [3]
RP GENE STRUCTURE.
RX MEDLINE=91161629; Pubmed=2002063;
RA Lincke C.R., Smit J.J.M., van der Velde-Koerts T., Borst P.;
RT "Structure of the human MDR3 gene and physical mapping of the human
RT MDR locus.";
RL J. Biol. Chem. 266:5303-5310(1991).
CC -!- FUNCTION: ENERGY-DEPENDENT EFFLUX PUMP RESPONSIBLE FOR DECREASED
CC DRUG ACCUMULATION IN MULTIDRUG-RESISTANT CELLS. HUMAN MDR3 IS NOT
CC OF PHOSPHATIDYLCHOLINE ACROSS THE CANALICULAR MEMBRANE OF THE
CC HEPATOCYTE.
CC -!- SUBCELLULAR LOCATION: Integral membrane protein.
CC -!- DISEASE: DEFECTS IN ABCB4 ARE A CAUSE OF PROGRESSIVE FAMILIAL
CC INTRAHEPATIC CHOLESTASIS TYPE III (PFIC), A FORM OF AUTOSOMAL
CC RECESSIVE LIVER DISORDERS, CHARACTERIZED BY EARLY ONSET OF
CC CHOLESTASIS THAT PROGRESSES TO CIRRHOSIS AND LIVER FAILURE BEFORE
CC ADULTHOOD.
CC -!- SIMILARITY: BELONGS TO THE ABC TRANSPORTER FAMILY. MDR SUBFAMILY.
CC
CC THIS SWISS-PROT entry is copyrighted. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@sib-sib.ch).
CC
DR EMBL: M23234; AAA36207.1; -.
DR EMBL: X06181; CAA29547.1; -.
DR PIR: J50051; DVH03.
DR PIR: A42213; A42213.
DR HSSP: P13569; INBD.
DR GENE: HGNC:45; ABCB4.
DR MIM: 171060; -.
DR MIM: 602347; -.
DR InterPro: IPR003593; AAA_ATPase.
DR InterPro: IPR003439; ABC_transporter.
DR InterPro: IPR001140; ABCtransportrTM.
DR Pfam: PF00005; ABC_tran; 2.
DR Pfam: PF00664; ABC_membrane; 2.
DR Prodom: PD000006; ABC_transporter; 2.
DR SMART: SM00382; AAA; 2.
DR PROSITE: PS00211; ABC_TRANSPORTER; 2.
KW ATP-binding; glycoprotein; Transmembrane; Transport; Repeat;
KW Multigene family.
FT DOMAIN 1 57 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 58 78 POTENTIAL.
FT TRANSMEM 123 143 POTENTIAL.
FT TRANSMEM 192 211 POTENTIAL.
FT TRANSMEM 216 235 POTENTIAL.
FT TRANSMEM 301 320 POTENTIAL.
FT TRANSMEM 336 356 POTENTIAL.
FT DOMAIN 355 711 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 712 732 POTENTIAL.
FT TRANSMEM 756 776 POTENTIAL.
FT TRANSMEM 832 851 POTENTIAL.

```

```

FT TRANSMEM 854 873 POTENTIAL.
FT TRANSMEM 937 956 POTENTIAL.
FT TRANSMEM 976 993 POTENTIAL.
FT TRANSMEM 994 1279 CYTOPLASMIC (POTENTIAL).
FT NP_BIND 429 436 ATP (BY SIMILARITY).
FT NP_BIND 1069 1076 ATP (BY SIMILARITY).
FT REPEAT 1 640
FT REPEAT 641 1279
FT CARBOHYD 91 91 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 97 97 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CONFLICT 1093 1093 V -> VEFDFGQ (IN REF. 2).
FT SEQUENCE 1279 AA: 140682 MW: 3058C98B5C8D6087 CRC64;

Query Match 75.7%; Score 4905.5; DB 1; Length 1279;
Best Local Similarity 75.5%; Pred. No. 1.2e-269;
Matches 973; Conservative 131; Mismatches 165; Indels 19; Gaps 9;

OY 1 MDPEGRKG-----SAKNEWMKMKSKRNEKKEK-PTVSTFPMFRYSNWIDRLIMLY 53
DB 1 MDLEAKNGCTAMRPSAEGDE-ELGISSSKOKRKKTVMKIGVLTFLFRYSMDQDKLMSL 59
OY 54 GTMAIIGHGALPLMVLFGNMTDSFANAGISRKTEPPVILNFINNTQHFINHLEEM 113
DB 60 GTIMAIAGSGIPLMMIVFGEMTDKFDVT--AGNFSFVNFSLNLPGR-----ILEEM 113
OY 114 TTYAVYSGIGAVLVAAVYQVSEFCLAGROIIRKQFHHALMROEIGMFDVHDVDEL 173
DB 114 TTKAVYSGIGAGVLAAYVQVSEFCLAGROIIRKQFHHALMROEIGMFDVHDVDEL 173
OY 174 NTRLTDVSKINEIGDKIMFHSIATFPGTIVGFRGKLLVLILATSPVLGSAI 233
DB 174 NTRLTDVSKISEGIGDYGMFQAVATFAGTIVGFRGKLLVLILATSPVLGSAI 233
OY 234 VAKTIISFDEKLLVYAKAGAAVEVLAIATVIAFGQKLEBKRYNKMLEAKGICIK 293
DB 234 VAKTIISAFDEKLLVYAKAGAAVEVLAIATVIAFGQKLEBKRYNKMLEAKGICIK 293
OY 294 AITANISIGAPLLIYASYALAFWTGTSIYASSEISQIVLTPSPVILIGAFSGSPS 353
DB 294 AITANISIGAPLLIYASYALAFWTGTSIYASSEISQIVLTPSPVILIGAFSGSPS 353
OY 354 IEFANARAGAAVEIFKIDNKPISDYSKSGHPNINIKNEFFKNVHSYPSRREVKILK 413
DB 354 IEFANARAGAAVEIFKIDNKPISDYSKSGHPNINIKNEFFKNVHSYPSRREVKILK 413
OY 414 GLNLKYSGGTVALVGNSSGKSTVQLMQLYDPLDGMVCLDGDITINVRHLREITG 473
DB 414 GLNLKYSGGTVALVGNSSGKSTVQLMQLYDPLDGMVCLDGDITINVRHLREITG 473
OY 474 VVSQEPVLEFNTTAENIRYGRNVTMDELEKAYKEANAYDFIMKLPNKFDLVGERGAOL 533
DB 474 VVSQEPVLEFNTTAENIRYGRNVTMDELEKAYKEANAYDFIMKLPNKFDLVGERGAOL 533
OY 534 SGGOKRIATARAIVRNPKILLDEATSAIDTESAEVVOVALDKARKGTTIVIAHRLST 593
DB 534 SGGOKRIATARAIVRNPKILLDEATSAIDTESAEVVOVALDKARKGTTIVIAHRLST 593
OY 594 VRNADVLAGFDGVIYVKGNDLMKKEGTYEKLVTMORGNIELENTYGESKESDAL 653
DB 594 VRNADVLAGFDGVIYVKGNDLMKKEGTYEKLVTMORGNIELENTYGESKESDAL 653
OY 654 EMSRKDSSSLIKRSTRRSIIHAPOGDRKIGTKED-LMENVPVPSFWRLKINSTEMPY 712
DB 654 EMSRKDSSSLIKRSTRRSIIHAPOGDRKIGTKED-LMENVPVPSFWRLKINSTEMPY 712
OY 713 FVVGIFCAIINGIOPAFSIIISHRIIGIFTRDEDEPFTRKONSNMFSVLVIGIISFIF 772
DB 713 FVVGIFCAIINGIOPAFSIIISHRIIGIFTRDEDEPFTRKONSNMFSVLVIGIISFIF 772
OY 773 FVVGIFCAIINGIOPAFSIIISHRIIGIFTRDEDEPFTRKONSNMFSVLVIGIISFIF 772
DB 773 FVVGIFCAIINGIOPAFSIIISHRIIGIFTRDEDEPFTRKONSNMFSVLVIGIISFIF 772

```

```

OY 833 SRLAVITONIANIGTIIISLYGQWQLLLLAIVPIIAIAGVEMKMLSGOALKRREL 892
DB 830 TRLALNOMIANIGTIIISLYGQWQLLLLAIVPIIAIAGVEMKMLSGOALKRREL 889
OY 893 EGAGKITEAIEENFRIVSLTREOKFEYMYAOSLOVPIYNSLRKAHIGVFSFTOMMY 952
DB 890 EAAGKITEAIEENFRIVSLTREOKFEYMYAOSLOVPIYNSLRKAHIGVFSFTOMMY 949
OY 953 FSYAGCFRCATVYANDEFANFODVILVESAIVFGAMAVGVSEFADYAKKVAHVIM 1012
DB 950 FSYAGCFRCATVYANDEFANFODVILVESAIVFGAMAVGVSEFADYAKKVAHVIM 1009
OY 1013 IIEKSPILDSYSPHGLKPNTEGNTFNEVENVYPTRPDIPLYGLGSLSEYKKGQTLAVG 1072
DB 1010 LFEKQPLIDSYSEGLKPKDFEGNITFNEVENVYPTRPDIPLYGLGSLSEYKKGQTLAVG 1069
OY 1073 SSGGKSTVQVLLERFYPDLAGSVLDGKEIKILANQWLRHNGIYSEPIIFDPSIAEN 1132
DB 1070 SSGGKSTVQVLLERFYPDLAGSVLDGKEIKILANQWLRHNGIYSEPIIFDPSIAEN 1129
OY 1133 IAYGDNRSVYSHETIMQAKEANIHPIETLPKRYNTRVGDGKTQLSGOKORAIARAL 1192
DB 1130 IAYGDNRSVYSHETIMQAKEANIHPIETLPKRYNTRVGDGKTQLSGOKORAIARAL 1189
OY 1193 VROPHILLDEATSAIDTESKVVQALDKARREGRCITVIAHRLSTIONADIVFONCR 1252
DB 1190 IROPHILLDEATSAIDTESKVVQALDKARREGRCITVIAHRLSTIONADIVFONCR 1249
OY 1253 VKEHGHOOOLLAOKGIYFSWVVOAGAK 1280
DB 1250 VKEHGHOOOLLAOKGIYFSWVVOAGAK 1277

RESULT 8
MDR2_MOUSE STANDARD; PRT; 1276 AA.
ID MDR2_MOUSE
AC P21440.1991 (Rel. 18, Created)
DT 01-MAY-1991 (Rel. 18, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Multidrug resistance protein 2 (p-glycoprotein 2).
GN ABCB4 OR PGY2 OR PGY-2 OR MDR2.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OC NCB1_TaxID=10090;
OX [1]
RN SEQUENCE FROM N.A.
RP MEDLINE=88302195; PubMed=3405218;
RX Gros P., Raymond M., Bell J., Housman D.;
RA "Cloning and characterization of a second member of the mouse mdr
RT gene family."
RL Mol. Cell. Biol. 8:2770-2778(1988).
RN [2]
RN SEQUENCE OF 1-23 FROM N.A.
RC SPRAIN-BALB/C;
RA Kirschner L.S., Horwitz S.B.;
RL Submitted (DEC-1991) to the EMBL/Genbank/DBJ databases.
CC -1- FUNCTION: ENERGY-DEPENDENT EFFLUX PUMP RESPONSIBLE FOR DECREASED
CC DRUG ACCUMULATION IN MULTIDRUG-RESISTANT CELLS. MOUSE MDR2 IS
CC NOT CAPABLE OF CONFERRING DRUG RESISTANCE.
CC -1- SUBCELLULAR LOCATION: Integral membrane protein.
CC -1- MISCELLANEOUS: IN MOUSE THE MDR GENE FAMILY INCLUDES THREE OR MORE
CC RELATED BUT DISTINCT CELLULAR GENES.
CC -1- SIMILARITY: BELONGS TO THE ABC TRANSPORTER FAMILY. MDR SUBFAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sdb.ch/announce/
CC or send an email to license@isb-sdb.ch).

```


Query Match	75.0%;	Score 4856;	DB 1;	Length 1276;
Best Local Similarity	74.7%;	Pred. No. 7.8e-267;		
Matches 959;	Conservative 139;	Mismatches 172;	Indels 14;	Gaps 8;

[illegible]

RX MEDLINE=92135896; PubMed=1685679;
RA Endicott J.A., Sarangi F., Ling V.;
RT "Complete cDNA sequences encoding the Chinese hamster P-glycoprotein
gene family.";
RL DNA Seq. 2:89-101(1991).
CC -1- FUNCTION: ENERGY-DEPENDENT EFFLUX PUMP RESPONSIBLE FOR DECREASED
CC DRUG ACCUMULATION IN MULTIDRUG-RESISTANT CELLS.
CC -1- SUBCELLULAR LOCATION: Integral membrane protein.
CC -1- MISCELLANEOUS: PGP ISOFORMS DIFFER IN THEIR DRUG TRANSPORT
CC CAPABILITIES: PGP1 AND PGP2 CAN MEDIATE MDR, WHILE PGP3 APPARENTLY
CC CANNOT.
CC -1- SIMILARITY: BELONGS TO THE ABC TRANSPORTER FAMILY. MDR SUBFAMILY.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: M60042; AAA68885.1; -.
DR HSSP: P13569; 1NBD.
DR InterPro: IPR003593; AAA_ATPase.
DR InterPro: IPR003439; ABC_transportr.
DR InterPro: IPR001140; ABCtransprtTM.
DR Pfam: PF00005; ABC_tran; 2.
DR Pfam: PF00664; ABC_membrane; 2.
DR ProDom: PD000006; ABC_transportr; 2.
DR SMART: SM00382; AAA; 2.
DR PROSITE: PS00211; ABC_TRANSPORTER; 1.
DR ATP-binding; Glycoprotein; Transmembrane; Transport; Repeat;
KW Multigene family.
FT DOMAIN 1 57 CYTOPLASMIC (POTENTIAL).
FT TRANSSEM 58 78 POTENTIAL.
FT TRANSSEM 122 142 POTENTIAL.
FT TRANSSEM 191 211 POTENTIAL.
FT TRANSSEM 218 238 POTENTIAL.
FT TRANSSEM 299 319 POTENTIAL.
FT TRANSSEM 328 348 POTENTIAL.
FT DOMAIN 349 712 CYTOPLASMIC (POTENTIAL).
FT TRANSSEM 713 733 POTENTIAL.
FT TRANSSEM 758 778 POTENTIAL.
FT TRANSSEM 834 854 POTENTIAL.
FT TRANSSEM 855 875 POTENTIAL.
FT TRANSSEM 938 958 POTENTIAL.
FT TRANSSEM 975 995 POTENTIAL.
FT TRANSSEM 996 1281 CYTOPLASMIC (POTENTIAL).
FT NP_BIND 429 436 ATP (POTENTIAL).
FT NP_BIND 1071 1078 ATP (POTENTIAL).
SO SEQUENCE 1281 AA; 140866 MW; 2203EF61BB29602 CRC64;

Query Match 74.4%; Score 4818.5; DB 1; Length 1281;
Best Local Similarity 73.6%; Pred. No. 1e-264;
Matches 948; Conservative 147; Mismatches 176; Indels 17; Gaps 8;

OY 1 MDPEGRGSAEK-----NFMKMGKSKKNKKKKKKPT-VSTAMFYSWMLRLMLVG 54
DB 1 MDLEAARNGTARRPGTVEGDELISISOGNKKKYNLGLPLTFPYSWODKLPMLG 60
OY 55 TMAAIIHGALPLMLVGNMTDSFA-NAGISRNKTFPVIIINESITNTTOHFINHEEM 113
DB 61 TTMALAHSGGLPLMIVGEMTDKFNAG---NFSLPVNSLSMIMGK-----ILEEEM 113
OY 114 TTYAAYYSIGAGLVAAVIOVSFWCLAGQOLIKRKOPFHAIIMROEIGWVHYDEL 173
DB 114 TTYAAYYSIGAGLVAAVIOVSFWCLAGQOLIKRKOPFHAIIMROEIGWVHYDEL 173
OY 174 NTRLDDVSKINIEGDKIGMFHSHIAFTFGTIVGFGRKLLVLVLATISPVLGSAI 233
DB 174 NTRLDDVSKINIEGDKIGMFHSHIAFTFGTIVGFGRKLLVLVLATISPVLGSAI 233
OY 234 WAKIISFDRKELAVAKAGAAVEVLAIKRTVIAFGQKKELRYNKNLEAKGIGIKK 293
DB 234 WAKIISFDRKELAVAKAGAAVEVLAIKRTVIAFGQKKELRYNKNLEAKGIGIKK 293

DB 234 WAKIISFDRKELAVAKAGAAVEVLAIKRTVIAFGQKKELRYNKNLEAKGIGIKK 293
OY 294 AITANISGAFLIYASALAEVYGSIVLSSEYSGOVLVTFVFSVLIAFSIGASPS 353
DB 294 AITANISGAFLIYASALAEVYGSIVLSSEYSGOVLVTFVFSVLIAFSIGASPS 353
OY 354 IAFANARGAATFIRKIDNKPISIDYSKSGKPKDNKGNLEFNHFSYPSRREVKILK 413
DB 354 IAFANARGAATFIRKIDNKPISIDYSKSGKPKDNKGNLEFNHFSYPSRREVKILK 413
OY 414 GNLKVGSGQVYALVNGSGCGKSTVQLMQRVDPDGMVCIDQDIRTINVRHLRITG 473
DB 414 GNLKVGSGQVYALVNGSGCGKSTVQLMQRVDPDGMVCIDQDIRTINVRHLRITG 473
OY 474 VVSQEPVLFATTIAENIRGRNVTMDLEKAVKANAYDIFMKLPKPFVLVGERGAOL 533
DB 474 VVSQEPVLFATTIAENIRGRNVTMDLEKAVKANAYDIFMKLPKPFVLVGERGAOL 533
OY 534 SGQOKORIAARLVNPKITLLDEATSLDTESEAVVQALDKARKGRTTIVIAHRLST 593
DB 534 SGQOKORIAARLVNPKITLLDEATSLDTESEAVVQALDKARKGRTTIVIAHRLST 593
OY 594 VRNADYIAGFDGVIYKGNHDELMKEGIRFKLYTQTRNEIELENAATESSESOL 653
DB 594 VRNADYIAGFDGVIYKGNHDELMKEGIRFKLYTQTRNEIELENAATESSESOL 653
OY 654 EMSPKSGSLIKRSTRSHAPQODRLGTAKD-LNENVPVSEFRIILKLNSTEMP 712
DB 654 GMP-NGMWSHIFRNSTKSLKSSRAHHRLVDVDELDAWNPVPSFLKVLKLNTEMP 712
OY 713 FVVGICAILNGIOPAFSITSRIGITRDEDPETKRONSMMFSVFLVGLIISFTF 772
DB 713 FVVGICAILNGIOPAFSITSRIGITRDEDPETKRONSMMFSVFLVGLIISFTF 772
OY 773 FLOGFTGKAGELITRKLRYMYFRSMLRODVSMPDCKMTGALTTRIANDAOVKGAIG 832
DB 773 FLOGFTGKAGELITRKLRYMYFRSMLRODVSMPDCKMTGALTTRIANDAOVKGAIG 832
OY 833 SRLAVIQTANTANGTGIISLIYGMQUTLLLAIPILAIAGVEMKMLSGOALKDKEL 892
DB 833 SRLAVIQTANTANGTGIISLIYGMQUTLLLAIPILAIAGVEMKMLSGOALKDKEL 892
OY 893 EGAKRIATEALENFRVVSILTRQKFEWTAQSLQVPRNSLRKAHIFCVSSTIOAMV 952
DB 893 EGAKRIATEALENFRVVSILTRQKFEWTAQSLQVPRNSLRKAHIFCVSSTIOAMV 952
OY 952 FSYAGCRRFGATVILVNGHMFEDVILVFSALYGAVALHASSFADYAKAKLSAHLFS 1011
DB 952 FSYAGCRRFGATVILVNGHMFEDVILVFSALYGAVALHASSFADYAKAKLSAHLFS 1011
OY 1013 IIEKSPILDSYSPHGLKPNLTLEGNVTENEVENYPTRPDIIVYLGISLEVKKGQTLALVG 1072
DB 1013 IIEKSPILDSYSPHGLKPNLTLEGNVTENEVENYPTRPDIIVYLGISLEVKKGQTLALVG 1072
OY 1072 SSGGCKSTVQVLLRFPDPLAGSVLIDGKEIKHLNVOMLRANHGIYSOPILFDCSIAEN 1132
DB 1072 SSGGCKSTVQVLLRFPDPLAGSVLIDGKEIKHLNVOMLRANHGIYSOPILFDCSIAEN 1132
OY 1133 IAYGDNSSVYVHEEIMOAAKEANITHFLETPPEKYNTRWGDKTQSGGOKORIAARLV 1192
DB 1133 IAYGDNSSVYVHEEIMOAAKEANITHFLETPPEKYNTRWGDKTQSGGOKORIAARLV 1192
OY 1193 VROPHTLLDEATSLDTESEKVVQELDKARRGRTCTIVIAHRLSTIQANADLVIVQNGK 1252
DB 1193 VROPHTLLDEATSLDTESEKVVQELDKARRGRTCTIVIAHRLSTIQANADLVIVQNGK 1252
OY 1252 VKEHGTQOOLAKGCIYFSMVNIOGAO 1279
DB 1252 VKEHGTQOOLAKGCIYFSMVNIOGAO 1279
RESULT 10

Db	61	ALANGSIPLMIMVIFGEMTDKEVDNAG---NFSLPVNESLMLNBR----ILREMBTRY	1131
Qy	117	AYVYSGAGVLAAYIOWSFMCIAAGROILKIROFPAHIREOIGFVDHVDGELNTR	176
Db	114	AYVYSGGVLAAAYIQVSWFTLAAGROIRIKOFKPHALIRQMGHFDIKGTTELNTR	173
Qy	177	LTDVSKINGIDGKIGMFHSHATFTGTFYGFTRGMKTLVILIASPVILGSAAIMV	236
Db	174	LTDISKISBIDGKMGFQOALTEPFAFIVGFTIRGMKTLVILIAITAIIGLSTAWAK	233
Qy	237	ILSFTRKELLAAYAKAGAAVEEVLAIPIVIAFGOKRELEBYNKLENAKGIKRAIT	296
Db	234	ILSFSDKELLAAYAKAGAAVEELGIRIVIAFGQNKLEBYOKHLENAKGIKRAIS	293
Qy	297	ANISGAFLIILYASTALAPWGTSLVISSEXYIGQVLTFFSVLIGAFSIGQASPIEA	356
Db	294	ANISMGAFLIILYASTALAPWGTSLVISSEXYIGQVLTFFSVLIGAFSIGQASPIDA	353
Qy	357	PANARGAAYEPIIDNKSIDSYSKSGKPNIKGNLEFNKMHFSPSKREVKYILGLN	416
Db	354	PANARGAAYEPIIDNKNKIDSFSBRGHKPUKIGNLEFSDVHSPYSRANKIILGLN	413
Qy	417	LKVOGQVYALVNGSCGKSTYVOLMORLYPTDQWVIDSQDIRTIVIRHLREITGV	476
Db	414	LKVSQGVYALVNGSCGKSTYVOLMORLYPTDQWVIDSQDIRTIVIRHLREITGV	473
Qy	477	QOEVLRTATIAEMIRGRNRYMDELKAYKANAAYPIKMLPKMPTLVGERAOISGC	536
Db	474	QOEVLRTATIAEMIRGRNRYMDELKAYKANAAYPIKMLPKMPTLVGERAOISGC	533
Qy	537	OKORIAIARALVNRPKILILDEATLALDTESEAVQVYALDKARGRTIYIARLSTVR	596
Db	534	OKORIAIARALVNRPKILILDEATLALDTESEAVQVYALDKARGRTIYIARLSTVR	593
Qy	597	ADVIAFPDQGVYVEKGNDELMEKGYIFKLYVMOTRGNELTELENATGESKSSDALMS	656
Db	594	ADVIAFEGGVYVEGSHBELKKGIFRVLNNOTSGOILSEEFLEVELDKAAGVA	653
Qy	657	PKDSGSLIKRRSTSRSHIAPOGQOQRIGTK-EDLENVVPVSWRLIKTNEMEYFV	715
Db	654	P-NGMKARIFERNSTKSLKSRSAHONRLDVTNELDANVPVSLKAKRLNKTEYFV	712
Qy	716	GIPCALINGLOPASFISPIRSIIGITFREDEDETRKONSNMFVFLVIGIISFTFFIQ	775
Db	713	GIPLCALANALOPAFSISPIRSIIGITFREDEDETRKONSNMFVFLVIGIISFTFFIQ	771
Qy	776	GFPPGAGGTTIKRLRYWVFPMSLRQDVSMDDPKNTGALTTRRLANDAQVGAIGSTL	835
Db	772	GFPPGAGGTTITRLRSWAFRAMLRODMSWMDHKNSTGALSTRLATDAOVOGATGRSL	831
Qy	836	AVITONIANLGTIIISLYCMQTLTLLLAIVPITALAGVEENKMLSGOALKRKELEBA	895
Db	832	ALIAONTANLGTIIISLYCMQTLTLLLSVPELIVAGVEENKMLAGNAKDKKEMEA	891
Qy	896	GKIAIEALENRTYVSLTREKFEYMYAQSLQVPYRNSLKRAHIFGVSWFSITQAMWFSY	955
Db	892	GKIAIEALENRTYVSLTREKFEYMYAQSLQVPYRNSLKRAHIFGVSWFSITQAMWFSY	951
Qy	956	AGCFEFGAYIVLANEFMNFODVLVAFSIVGGAANOVSFADYAKAKVSAAHYIMIE	1015
Db	952	AGCFEFGAYIVLANEFMNFODVLVAFSIVGGAANOVSFADYAKAKVSAAHYIMIE	1011
Qy	1016	KSPILDISPBGKLPNTLEGNVTFNEVENYPTRPDIPIVLOGLSLEVKKGQTLALVSSG	1075
Db	1012	ROPILDISYREGMMPDKFEESVTFNEVENYPTPRANVPYLOGISLEVKKGQTLALVSSG	1071
Qy	1076	CGKSTVVOLERFDPPLAGSLVDGKRIKLANQMLRAHGLIVSOEPIITDOSTANENAY	1135
Db	1072	CGKSTVVOLERFDPPLAGSLVDGQPAKKLANQMLRAOGLIVSOEPIITDOSTAKNATV	1131
Qy	1136	GDNSRVVSHETIQAAKEANIHHFIETLDEKVTNRGDDGTOLSGOKORIAIARALVQ	1195
Db	1132	GDNSRVVSOEIVIAKEANIHHFIETLDEKVTNRGDDGTOLSGOKORIAIARALVQ	1191

Oy	1196	PHILIDEATSAIDTSESEYVOEALDKRAREGTCIVIAHRLSTIONANDLIVYONSKYVE	1255
		I :	
Db	1192	PRVILIDEATSAIDTSESEYVOEALDKRAREGTCIVIAHRLSTIONANDLIVYONSKYVE	1251
Oy	1256	HGHQOLLAOKGIYFSMVSVOAGAK	1280
Db	1252	HGHQOLLAOKGIYFSMVINQAGTO	1276
RESULT 11			
ID	AB11_HUMAN	STANDARD;	PRT: 1321 AA.
AC	O95342; O9UNB2;		
DT	16-OCT-2001 (Rel. 40, Created)		
DT	16-OCT-2001 (Rel. 40, Last sequence update)		
DT	15-JUN-2002 (Rel. 41, Last annotation update)		
DE	Bile salt export pump (ATP-binding cassette, sub-family B, member 11).		
GN	ABCB11 OR BSEP.		
OS	Homo sapiens (Human).		
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		
OC	Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Homo.		
OX	NCBI_TaxID=9606;		
RN	[1]	SEQUENCE FROM N.A. AND VARIANTS PFIC2 GLY-297; GLU-461; GLY-482;	
RP	ARG-982; CYS-1153 AND GLN-1268.		
RP	MEDLINE=99021377; PubMed=9806540;		
RA	Stratuliets S.S., Bull L.N., Katsely A.S., Kocoshv S.A., Dahl N.,		
RA	Arneil H., Sokal E., Dahan K., Childs S., Ling V., Tanner M.S.,		
RA	Kaglawalla A.F., Nemech A., Pavlovskaja J., Baker A., Melli-Vergani G.,		
RA	Frelmer N.B., Gardiner R.M., Thompson R.J.;		
RT	"A gene encoding a liver-specific ABC transporter is mutated in		
RT	progressive familial intrahepatic cholestasis.";		
RL	Nat. Genet. 20:233-238(1998).		
RN	[2]	SEQUENCE FROM N.A.	
RA	Mol O., Hooiveld G.J.E.J., Jansen P.T.M., Muller M.;		
RT	"Cellular localization and functional characterization of the human		
RT	bile salt export pump (BSEP).";		
RT	Submitted (MAR-1999) to the EMBL/Genbank/DBJ databases.		
RN	[3]		
RP	VARIANT ALA-444.		
RX	MEDLINE=21686803; PubMed=11829140;		
RA	Saito S., Iida A., Sekine A., Miura Y., Ogawa C., Kawachi S.,		
RA	Higuchi S., Nakamura Y.;		
RT	"Three hundred twenty-six genetic variations in genes encoding nine		
RT	members of ATP-binding cassette, subfamily B (ABCB/MDR/TAP), in the		
RT	Japanese population.";		
RL	J. Hum. Genet. 47:38-50(2002).		
RN	[4]		
RP	VARIANTS INTRAHEPATIC CHOLESTASIS LEU-284 AND ASP-1004.		
RX	MEDLINE=21674594; PubMed=11815775;		
RA	Chen H.-L., Chang P.-S., Hsu H.-C., Ni Y.-H., Hsu H.-Y., Lee J.-H.,		
RA	Jeng Y.-M., Shau W.-Y., Chang M.-H.;		
RT	"FIC1 and BSEP defects in Taiwanese patients with chronic intrahepatic		
RT	cholestasis with low gamma-glutamyltranspeptidase levels.";		
RL	J. Pediatr. 140:119-124(2002).		
CC	-1- FUNCTION: INVOLVED IN THE ATP-DEPENDENT SECRETION OF BILE SALTS		
CC	INTO THE CANALICULUS OF HEPATOCYTES.		
CC	-1- SUBCELLULAR LOCATION: Integral membrane protein.		
CC	-1- TISSUE SPECIFICITY: EXPRESSED PREDOMINANTLY, IF NOT EXCLUSIVELY IN		
CC	THE LIVER, WHERE IT WAS FURTHER LOCALIZED TO THE CANALICULAR		
CC	MICROVILLI AND TO SUBCANALICULAR VESICLES OF THE HEPATOCYTES BY		
CC	SITE (BY SIMILARITY).		
CC	-1- DOMAIN: MULTIFUNCTIONAL POLYPEPTIDE WITH TWO HOMOLOGOUS HALVES,		
CC	EACH CONTAINING AN HYDROPHOBIC MEMBRANE-ANCHORING DOMAIN AND AN		
CC	ATP BINDING CASSETTE (ABC) DOMAIN.		
CC	-1- DISEASE: DEFECTS IN ABCB11 ARE THE CAUSE OF PROGRESSIVE FAMILIAL		
CC	INTRAHEPATIC CHOLESTASIS 2 (PFIC2), AN INHERITED LIVER DISEASE OF		
CC	CHILDHOOD. PFIC2 IS CHARACTERIZED BY CHOLESTASIS AND NORMAL SERUM		
CC	GAMMA-GUTAMINYLTRANSFERASE ACTIVITY. DEFECTS IN ABCB11 ARE ALSO		
CC	FOUND IN CASES OF CHRONIC INTRAHEPATIC CHOLESTASIS WITHOUT OBVIOUS		
CC	FAMILIAL HISTORY OF CHRONIC LIVER DISEASE.		

CC	-1- SIMILARITY: BELONGS TO THE ABC TRANSPORTER FAMILY. MOR SUBFAMILY.
CC	-----
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration
CC	between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC	the European Bioinformatics Institute. There are no restrictions on its
CC	use by non-profit institutions as long as its content is in no way
CC	modified and this statement is not removed. Usage by and for commercial
CC	entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC	or send an email to license@sib-sib.ch).
CC	-----
DR	EMBL: AF091582; AAC7455.1; -
DR	EMBL: AF136523; AAD28285.1; -
DR	Genew: HGNC:42; ABCB11.
DR	MIM: 603201; -
DR	MIM: 601847; -
DR	InterPro: IPR003593; AAA_ATPase.
DR	InterPro: IPR003439; ABC_transportr.
DR	InterPro: IPR01140; ABCtransportrTM.
DR	Pfam: PF00005; ABC_tran; 2.
DR	Pfam: PF00664; ABC_membrane; 2.
DR	ProDom: PD000006; ABC_transportr; 2.
DR	SMART; SM00382; AAA; 2.
DR	PROSITE; PS00211; ABC_TRANSPORTER; 1.
KW	ATP-binding; Transmembrane; Transport; Disease mutation; Polymorphism.
FT	DOMAIN 1 62
FT	TRANSMEM 63 83
FT	POTENTIAL POTENTIAL.
FT	DOMAIN 84 147
FT	TRANSMEM 148 168
FT	POTENTIAL POTENTIAL.
FT	DOMAIN 169 235
FT	TRANSMEM 216 236
FT	POTENTIAL POTENTIAL.
FT	DOMAIN 237 240
FT	TRANSMEM 241 261
FT	POTENTIAL POTENTIAL.
FT	DOMAIN 262 319
FT	TRANSMEM 320 340
FT	POTENTIAL POTENTIAL.
FT	DOMAIN 341 353
FT	TRANSMEM 354 374
FT	POTENTIAL POTENTIAL.
FT	DOMAIN 375 755
FT	TRANSMEM 756 776
FT	POTENTIAL POTENTIAL.
FT	DOMAIN 777 794
FT	TRANSMEM 795 815
FT	POTENTIAL POTENTIAL.
FT	DOMAIN 816 869
FT	TRANSMEM 870 890
FT	POTENTIAL POTENTIAL.
FT	TRANSMEM 891 911
FT	POTENTIAL POTENTIAL.
FT	DOMAIN 912 979
FT	TRANSMEM 980 1000
FT	POTENTIAL POTENTIAL.
FT	DOMAIN 1001 1011
FT	TRANSMEM 1012 1032
FT	POTENTIAL POTENTIAL.
FT	DOMAIN 1033 1321
FT	TRANSMEM 1033 1321
FT	POTENTIAL POTENTIAL.
FT	NP_BIND 455 462
FT	NP_BIND 1113 1120
FT	CARBOHYD 109 109
FT	CARBOHYD 116 116
FT	CARBOHYD 122 122
FT	CARBOHYD 125 125
FT	POTENTIAL POTENTIAL.
FT	VARIANT 284 284
FT	CHOLESTASIS).
FT	/FtId=VAR_013332.
FT	E -> G (IN PFIC2).
FT	/FtId=VAR_010271.
FT	V -> A.
FT	/FtId=VAR_013333.
FT	K -> E (IN PFIC2).
FT	/FtId=VAR_013334.
FT	D -> G (IN PFIC2).
FT	/FtId=VAR_013335.
FT	G -> R (IN PFIC2).
FT	/FtId=VAR_013336.
FT	G -> D (IN CHRONIC INTRAHEPATIC
FT	CHOLESTASIS).
FT	/FtId=VAR_013337.
FT	R -> C (IN PFIC2).
FT	/FtId=VAR_013338.
FT	R -> Q (IN PFIC2).
FT	VARIANT 1268 1268

```

FM          339          339          /FTID=VAR_013339.
FT          50          1321 AA: 146392 MM: D44ACACC48DEA371 CRC64:
SEQUENCE

Query Match      50.9%; Score 3300; DB 1; Length 1321;
Best Local Similarity 50.3%; Pred. No. 7,8e-179;
Matches 657; Conservative 242; Mismatches 367; Indels 40; Gaps 10;

QY      8 KGAENKFWMGKRSKKNKK-----EKK---PTVSTFAMFYSNMIDRLVMTVMTAI 59
      11 KKFGEENDGESDKSYNNDKSKSLDDEKKGDDVRRVGFQDLFFSSSTDIWIMLFSCAF 70
      60 IHGAALPLMALVYGNMTDSFANAGISR-----NKTFF---PVINESITNTQHF 105
      71 LHGIADPGVLLFGTWTDFIDYDVELDELQIPGKACVNMNTIVMNSLSNMNTNGTCG 130
QY      106 INLEEMFTTAAVYVYAGVLAAYVQVSWCLAAAGROLIKRKPFNAIMROETGWF 165
      131 LNTIESEMKEFASYTAGIYAVLITGYIQICFWIAAARQOKMKKFRRIMHEIGWF 190
QY      166 DVHVDGLNTRLLDVSKEINIGDKIGMFHSHIATFFGTGIVGTGKMTLVILASP 225
      191 DENSVEELMTRESDDINKINDADQMALFIQRMNSTICGFLDFPRGKMLTVIISVP 250
QY      226 VIGLSAIAIAKILTSFTDEKELATYAKAGVAEVLAIKIVIAFGGKRELERYNKNLEE 285
      251 LIGIGATIGLSVSKFTYELKAAKAGVADDEVISSMRYVAAGGKREVEREKNLVF 310
QY      286 AAGIGIKRAITANISIGAFLIYASVALAFWCGSLVL--SSEKISQVLTFFSVLGA 344
      311 AQRMGIRGIVYVGFPTGKWLIFLCYAVATWYSTVLDEGETTPTVQILSYLVGA 370
QY      345 PSIGASISIAFANAGAAVEIFKIIDNKPSTDSYKSGHKPDNIKNGLEFNKVFHSP 404
      371 LNLGASPCLEAFATGRAAATSIETIDRKPIDCMSEDEYKIDRIKGEIEFNVFHP 430
QY      405 SRKEVYKILKNTLVKOSQOTVALVNSGCKSTVOQLQMLYPTDGMVTCIDGIRTN 464
      431 SRPEKILINDMNTIKFGEEMALVGPAGAGSTALDLOIFRYPDCRMVVDHDIRSN 490
QY      465 VRLHREITGVVSOBPLVATTIAINIRYGRENVYMDIEKAVEKANAADPEIMLPKPF 524
      491 IOWLRDQIGIYEOBPVLESTIAENIRYGRBDATMEDIVQAEEANAYNFIMLPQOFT 550
QY      525 LVYRGQOLSGGKORITAIARALYRNKILLLDEATSDLTSEAVVOYALDKARKRTT 584
      551 LVYEGGOMSGGOKORVAIARALIRNPKILLDMATSDLDNESEAAVQEVLSKHGHTI 610
QY      565 IVIAHRLSTVANADVIAGFDGVIVKGNHDELMKEGIFYFKLYMTQGS-----NEIELE 640
      611 ISVAHRLSTVRAADTIIGFEHGTAVEGHTEILERKGYFTLVLTLOSOGNOLNEEDIK 670
QY      641 NATPESESSEDALEMPKDSGLIKRRSTR--SIHAP--OGODRKIGTGED----- 689
      671 DATDDMLARTFSGSDSLRASIRORSKQSLSYLVHPEPLAVVDHKTSTYEDKMDI 730
QY      690 -LANNVPPVGFMRILKLNSTEMPEYFVNGIFCAIINGLOPAPSIIFSRITIGFTDEDE 748
      731 PVGEVEPAPVRRILKFSAPEMPYMLVSGAVAVNGTVPPLAFLFSLQITGFS-IPKPE 789
QY      749 TKRONSMESVFLVLIISITFFLQGTFFGACGIELTKRLRYVFRSMLEQDVSFDP 808
      790 -EORSQINGVCLLVFVAMCVSLFQFQGYAFKSGSEGLLTKRKKGFRAMLGODIAMDD 849
QY      809 PKNTGTALTIRLANDAAOYKGAIGSLAVIYTONIANLGTGIIISLYQMQLTLLLATVP 868
      850 LRNSPQALTRLATDSQVOGAGSOGIMIVSFNVMIAIARSFPMKSLVLTICFP 909
QY      869 IAIAGVEMKMLSGQALDKKELEGAKRIATEENFRVYSLRREKFEYMAQSLQV 928
      910 FLALSGATQTRMLTGFRASDKQALEWVGQITTEALSNTKRYAGIKERKRFLEALETELEK 969
QY      929 PYNSLRKRIHIGVSPSITQAMMYFSYAGCERFEGAYLVANEFMFDVLLVSAIVEGAM 988

```

```

      970 PFTALQKANIYGFCEFAQOCINFANSASYRGGLISNEGILHFSYFVIAVLSAT 1029
QY      989 AVGVSSFAPDYAKAKVYSAAHVIMIEKSPLDISPHOLKNTLEGVNTRENVFNTPT 1048
      1030 ALBRATSYTPSYAKAKISARFFQULDRQPPISVNTNGEKKMDNQGKIDVDCKFYPS 1089
QY      1049 RPDIPVLOQISLEVRKQGTALVVGSGGKSTVVOLLERFYDPLAGSLVIGKEIKHNV 1108
      1090 RPDQVNLNGISVISPQGLTAVVGSGGKSTISLLERFPDPDGKMAIDGHSKAYNV 1149
QY      1109 QMLRAHLGVSOEPILEFDSIAENTAYGDNKRVSHHEIMQAAEANIHFEITLPEKYN 1168
      1150 QFLRSNIGVISOEPIVAFCSINDNIKYGNTKEIPMERVIAAAQOQLHDFVMSLPEKE 1209
QY      1169 TRYGDKGTOLSGGOKORIAIARALVROPHILLDEATSDLTSEKRYQVGLDKARKGRT 1228
      1210 TVNGSGGSLSGEKORIAIARALVROPHILLDEATSDLTSEKRYQVGLDKARKGRT 1269
QY      1229 CIVIAHRLSTIONADLIIVFPONGKVEKHTHQOQLAKGIFYSMWS 1274
      1270 CIVIAHRLSTIONADLIIVMAQGVIEKGTHEIMQAQGVAYKLVLT 1315

```

RESULT 12

```

AB11_RABIT
ID AB11_RABIT STANDARD; PRT; 1321 AA.
AC 09NOV03;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Bile salt export pump (ATP-binding cassette, sub-family B, member 11)
DE (Sister of P-glycoprotein).
GN ABCB11 OR BSEP OR SPGP.
OS Oryctolagus cuniculus (Rabbit).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OM Mammalia; Eutheria; Lagomorpha; Leporidae; Oryctolagus.
OX NCBI_TaxId=9986;
RN [1]
RP
RA Balasubramanian N.V., Suchy F.J., Ananthanarayanan M.;
RT "Molecular cloning and characterization of rabbit liver bile salt
  export pump (Bsep/spgp).";
RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.
CC
CC -1- FUNCTION: INVOLVED IN THE ATP-DEPENDENT SECRETION OF BILE SALTS
  INTO THE CANALICULUS OF HEPATOCYTES.
CC
CC -1- SUBCELLULAR LOCATION: Integral membrane protein.
CC
CC -1- TISSUE SPECIFICITY: EXPRESSED PREDOMINANTLY, IF NOT EXCLUSIVELY IN
  THE LIVER, WHERE IT WAS FURTHER LOCALIZED TO THE CANALICULAR
  MICROVILLI AND TO SUBCANALICULAR VESICLES OF THE HEPATOCYTES BY IN
  SITU.
CC
CC -1- DOMAIN: MULTIFUNCTIONAL POLYPEPTIDE WITH TWO HOMOLOGOUS HALVES,
  EACH CONTAINING AN HYDROPHOBIC MEMBRANE-ANCHORING DOMAIN AND AN
  ATP BINDING CASSETTE (ABC) DOMAIN.
CC
CC -1- SIMILARITY: BELONGS TO THE ABC TRANSPORTER FAMILY. MDR SUBFAMILY.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
  between the Swiss Institute of Bioinformatics and the EMBL Outstation -
  the European Bioinformatics Institute. There are no restrictions on its
  use by non-profit institutions as long as its content is in no way
  modified and this statement is not removed. Usage by and for commercial
  entities requires a license agreement (see http://www.ebi.ac.uk/announcements/
  or send an email to license@ebi.ac.uk).
CC
CC EMBL: AF249879; AAF65552.1;
CC InterPro: IPR003593; AAA_Atpase.
CC InterPro: IPR003439; ABC_transporter.
CC InterPro: IPR001140; ABCtransportTM.
CC Pfam: PF00005; ABC_tran. 2.
CC Pfam: PF00664; ABC_membrane. 2.
CC ProDom: PD000006; ABC_transportr. 2.
CC SMART: SM00382; AAA; 2.

```

```

0Y 544 ARAIVRNPKILLDEATSALDTESAEVAOVALDKARKRRTTIVIAHLSTVYRNDADYAC 603
Db 570 ARAILIRNPKILLDMTASALDNESSEAMQOALSKTOHCHTIVSAHNRATIRTDADYIGC 629
0Y 604 DDGIVYKGNHMDLMEKNGIFYEFLVMOTRNGEIELE---NAVGESKSESDALEMSPKD 659
Db 630 EHGAEVGEDEEBLEELKNGYFALYALTOGQRNNGQOENEKMDATEDDIPKTFSGRYOD 689
0Y 660 SGSSLIKRRS-----TRSHAPOGQDRKGTKEDLNENPVPSFIRIL 703
Db 690 SLRSLRQRSKSOLSTYLAEHPMAVEDHKSTHEBRKDDLPDAOED---LEPSVYRIM 745
0Y 704 KLNSTNEMFYVYVGFCAIILINGLOPAFDSIFSRIGIIFTRDDEPPTKRONSNMSEVLELV 763
Db 746 KLNAPENMYTLLGSMGAANVGANTPLVALFESQILGTFSL-PDKREORSQINGICLLFVTV 804
0Y 764 LGIISFIFLFGTFESKAGBILTKRLIRWYFMSLRQOVSMFDDPKTKMTGATLIRLAND 823
Db 805 LGCVSEFFPOFGYTFPAKSGELLTKRLRKRFPAAMGODIGMFDDLRNSPGALTRLATID 864
0Y 824 LAQVNGALGSLAVITONINLGITISILTYEQMLTLLLAIVPIIAAGVEKMLSG 883
Db 865 ASQVYGAAGSGOIGMWNVSFTNIVYAMTIALFESKMLTLLIGVOPFFLALSGALQKMLTG 924
0Y 884 QALKDRKELEGAGKIALTEALENFTVVSLTREOKFETMYAQSLQVYRNSLRKAIIEVS 943
Db 925 PASRDKOALEKAGQITSEALSINRTVAGICKERKEFLETEAELEKRYKAIKANAVYOLC 984
0Y 944 FSTIQAMMYFYAGCFRGALVLANEPMNPODVLALFSAIVGAMNAVGVSSFAVDYAKA 1003
Db 985 FGPSCOTTFIANSASYRKGGVLLINBEGHLHSYFPRVYASAVLISATFALGRASSYTPSYAKA 1044
0Y 1004 KVSAAVHVIITEKSLIDSYSPHGKLRNTLEGVNTFHEVYENFTPTPDPIVYIGISLEVK 1063

```

RESULT 13		
AB11_MOUSE	STANDARD:	PRT: 1321 AA.
ID	AB11_MOUSE	
AC	Q90Y30; Q90ZE8;	
DT	16-OCT-2001 (Rel. 40, Created)	
DT	16-OCT-2001 (Rel. 40, Last sequence update)	
DT	16-OCT-2001 (Rel. 40, Last annotation update)	
DE	Bile salt export pump (ATP-binding cassette, sub-family B, member 11)	
DE	(Sister of P-glycoprotein).	
GN	ABCB11 OR BSEP OR SPCP.	
OS	Mus musculus (Mouse).	
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;	
OC	Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.	
OX	NCBI_Taxid=10090;	
RN	[1]	
RP	SEQUENCE FROM N.A.	
RC	TISSUE=Liver;	
RC	MEDLINE=20076398; PubMed=10607905;	
RA	Green R.M., Hoda F., Ward K.L.	

RT "Molecular cloning and characterization of the murine bile salt export pump.";
 RL Gene 241:117-123(2000).
 RN [2]
 RP SEQUENCE OF 463-635 FROM N.A.
 RC STRAIN-C57BL/6; TISSUE-Liver.
 RA Salkar R., Suchy F.J., Ananthanarayanan M.;
 RT "Molecular cloning of mouse liver bile salt export pump (bsep).";
 RL Submitted (SEP-1999) to the EMBL/Genbank/DBP databases.
 CC -1- FUNCTION: INVOLVED IN THE ATP-DEPENDENT SECRETION OF BILE SALTS INTO THE CANALICULUS OF HEPATOCYTES.
 CC -1- SUBCELLULAR LOCATION: Integral membrane protein.
 CC -1- TISSUE SPECIFICITY: EXPRESSED PREDOMINANTLY, IF NOT EXCLUSIVELY IN THE LIVER, WHERE IT WAS FURTHER LOCALIZED TO THE CANALICULAR MICROVILLI AND TO SUBCANALICULAR VESICLES OF THE HEPATOCYTES BY IN SITU.
 CC -1- DOMAIN: MULTIFUNCTIONAL POLYPEPTIDE WITH TWO HOMOLOGOUS HALVES, EACH CONTAINING AN HYDROPHOBIC MEMBRANE-ANCHORING DOMAIN AND AN -1- SIMILARITY: BELONGS TO THE ABC TRANSPORTER FAMILY. MDR SUBFAMILY.
 CC -----
 CC This SWISS-Prot entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL Outstation - use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL: AF133903; AAF14372.1; -
 DR EMBL: AF186585; AAD56419.1; -
 DR MGI: 1351619; Abcb11.
 DR InterPro: IPR003593; AAA_ATPase.
 DR InterPro: IPR003439; ABC_transport.
 DR Pfam: PF00005; ABC_tranp.fam.
 DR Pfam: PF00664; ABC_tranp.2.
 DR ProDom: PD000006; ABC_membrane.2.
 DR SMART: SM00382; AAA.2.
 DR PROSITE: PS00211; ABC_TRANSPORTER.1.
 DR KW ATP-binding; Transmembrane; Transport.
 FT DOMAIN 1 62
 FT TRANSMEM 63 83
 FT DOMAIN 84 147
 FT TRANSMEM 148 168
 FT DOMAIN 169 215
 FT TRANSMEM 216 236
 FT DOMAIN 237 240
 FT TRANSMEM 241 261
 FT DOMAIN 262 319
 FT TRANSMEM 320 340
 FT DOMAIN 341 353
 FT TRANSMEM 354 374
 FT TRANSMEM 375 755
 FT TRANSMEM 756 776
 FT TRANSMEM 777 794
 FT TRANSMEM 795 815
 FT TRANSMEM 816 869
 FT TRANSMEM 870 890
 FT TRANSMEM 891 911
 FT TRANSMEM 912 979
 FT TRANSMEM 980 1000
 FT TRANSMEM 1001 1011
 FT TRANSMEM 1012 1032
 FT TRANSMEM 1033 1321
 FT NP_BIND 455 462
 FT CARBOHYD 1113 1120
 FT CARBOHYD 109 109
 FT CARBOHYD 116 116
 FT CARBOHYD 122 116
 FT CARBOHYD 125 125
 FT CONFLICT 481 481
 FT CONFLICT 633 633
 FT CONFLICT 633 633

SQ SEQUENCE 1321 AA: 146675 MR: 1585EBF15D32967 CRC64;
 Query Match 49.5%; Score 3207.5; DR 1; Length 1321;
 Best Local Similarity 49.7%; Pred. No. 1,36-173;
 Matches 642; Conservative 237; Mismatches 377; Indels 37; Gaps 9;
 QY 15 FMKMGKSK-KNEKKKKPYSTFAMFRRYSNWLDRLYLVTMAATIGALPLMVLVFG 73
 Db 25 FHNNDKSKSLQDKKKKGGARVGFELFRFSSSKDNMTMFMGSCALDHGAQGMILIVG 84
 QY 74 NMIDSFANAGISR-----NKTFPVY---INESINNNQPHFINLHEEMTYANY 119
 Db 85 ILTDFEVEIEDIERQLSLIPKVCNNNTIWNINSEFNQNTNGISGIVDINSEVIRSGI 144
 QY 120 YSGIGAVLVAAIYOVSWICLAAGRLKTRKQFPAIMRQEGMDVDHVELMTRLFD 179
 Db 145 VAGVAVANLITGFIQIRLWMTTGARQIRKMKRYEFRIMMEIGWPDCTSVGELNSRSD 204
 QY 180 DVSKINIEGIDKIGMFHSIATFPTGFTVGFTGKWLITVITLAIAPVGLSAIYAKILS 239
 Db 205 DINKIDEAIDQMALFLQRLSTALSGILGFTYKMKLITLAIASVLIGIGAVIGLSVA 264
 QY 240 SETDKELIAYKAGAVAEVLAIRTYIAPGGQKELERYNNLEBANGIGIKKAITANI 299
 Db 265 KFTLELEKATYAKGSIADDEVLSIRTYAAGGKNEVEREYERKNLMFAQRWGIKGMVWG 324
 QY 300 SIGAFLITLYAVALAFWYGTSLV--SSEYSIGOVLTFFVSILGAFSIGQASSTIEAFA 358
 Db 325 FTGYMCLIFCYVALAFWYGSRLVLDGEYTGTLQITLCTIYIANNIMGNSSCLEIFS 384
 QY 359 NARGAAVEIKIIDNKPISDYSKSGHKPDNIKMLEKKNVHFSPSRKEVITKGLNLK 418
 Db 385 TCGSAASSIFQTDQPVMDCMGSDGQKIDRIKGEIEFHNVTYFHPSPREVKILNLSMV 444
 QY 419 VQSGQTVAAVGNSSCGKSTTVOLMQLYDPDGMVCIQDGOIRITNVRHLEITGVVSOE 478
 Db 445 IKGETTAFFVGSAGKSTALDLOIRFDPCGKWTTLGHDIRSLNLRMDQIGVIEOE 504
 QY 479 PVFATTLIENIRYRENTYMDIEIKAVKANAYDFIMKLPNKFDTLVGRGALSGSGOK 538
 Db 505 PVLESTTIENIRIGREEMEDIVQAAKDANAYNFIMALPQGDTLVGGGGMGSGOK 564
 QY 539 ORIAIRALVNRPKILLIDDEATSAIDTESEAVNOVALDKARKGTTTVIAHRLSTVRND 598
 Db 565 ORVALARALIRPKILLIDMATSAIDNESEAKVQALNKIQHGTIISVAHRLSTVRSAD 624
 QY 599 VTAGDDQIYVKGHDELMKEKGIYFKLVYMTQRGNEILENA--TGESKSESDALEM-- 655
 Db 625 VTIQEHGTAVRGTHIELIRKKGIVYEMLVYIQSDENHTKETIGSKQTDIEDPIERTE 684
 QY 656 ---SPKDGSSILIKRRSTR-----SIHAPQODRKLGTKEIDINENVPVSW 700
 Db 685 SRGVSQDSLRSINQKRSQSLSHSHEPPLAIDGHSYSYEDR--DNDAVLEVEVPAPVR 742
 QY 701 RLKLNSTENPFFVVGIFPCALINGLOPAFSIIFSRNIIIGTFRDEDPETKRNNSMPSVL 760
 Db 743 RLKYNISEMPYIIVGALCAINGAVPIYSLFQILITFGL-VDKEQQRSEIYSKCLF 801
 QY 761 FLVLGIISLIFFFLQFTFGKAGEIILRLRYVFSMSLRQVSWFDDPKNTGATLTRL 820
 Db 802 FVLIGCVSLTFLOPQGNFPAKSGELLTKRLRFGFAMLRQDIGNFDDLKNNPVLTRL 861
 QY 821 ANDAAQVKAIGSRRLAVTTONIANTLGTIIISLIYQWOTLLLALIVPILIAAGVETKM 880
 Db 862 ATDAISOVQATGSOVQVMMNSFNIFFVAVLIAFLFNWMLSLVIVSFFFLLSAQVOTKM 921
 QY 881 LSGQALKDKKELGAKIKNTALINERFVYSLTRBQKEVYMAQSLQVYNSLRKAHIF 940
 Db 922 LTGFPAQDKRILEKAGQITINEALSNIRYVAGIGVGRIFKAFVELEKSYKTAIRKANVY 991
 QY 941 GVSFSITQAMMYFSAVAGCFRFGAYLVANEFMNFQDVLVLPSAIYFGMAAGVQSSRPDY 1000
 Db 982 GLCYARSGISFIANSAAVRGYLIVYEDINFSIVRVVSSIAMSATAVAGRFSTYTPSY 1041

QY 1001 AAKYSAHVMITEKSPLDISYSPHGLKFNTEGWNFENYVNPTRPDIPVGLSLT 1060
 DB 1042 AAKKISAFAFPDLDRKRPDIIVYSGAGEKWNFGKIDICKFYPSPRPIQVNLNLSV 1101
 QY 1061 EYKSGQTALVSSGSGKSTVOLLERFYDPLAGSVLIDGKEIKHLNOMLRAHLGVISQ 1120
 DB 1102 SYDPGQTALFVSSGSGKSTISLLEFRFYDPDGTVMIDGHSKSKVNVQFLRSMIGIVSQ 1161
 QY 1121 EPLTPGCSIAENIANGDSRVSHHEEMOAKAEANHHEFLPEKTYRFGKSGQLSG 1180
 DB 1162 EPLVDPGCSLMDIKYGDWTKETISVERALPAKQOLHDSVMSLPKTKETNVIGISQSLR 1221
 QY 1181 GOKORIAIARALVROPHILLIDENATSDTESEKVVQENALDKAREGTCIVIAHRLSTIQ 1240
 DB 1222 GOKORIAIARALVROPHILLIDENATSDTESEKVVQENALDKAREGTCIVIAHRLSTIQ 1281
 QY 1241 NADLIYFQNGKVEHGHOOOLAOKGIFESWY 1273
 DB 1282 NSDIIAVMSQGVLEKGTBHKINDQKAYYKLV 1314

RESULT 14

AB11_RAT STANDARD; PRT: 1321 AA.
 ID AB11_RAT
 AC 070127;
 DT 16-OCT-2001 (Rel. 40, Created)
 DT 16-OCT-2001 (Rel. 40, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Bile salt export pump (ATP-binding cassette, sub-family B, member 11)
 DE (Sister of P-glycoprotein).
 GN ABCB11 OR ABSP OR SPGP
 OS Rattus norvegicus (Rat).
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
 ON NCBI_TaxId=10116;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=Sprague-Dawley; TISSUE=Liver;
 RX MEDLINE=98212048; PubMed=9545351;
 RA Gerloff T., Stieger B., Hagenbuch B., Madon J., Landmann L., Roth J.,
 RA Hofmann A.F., Meier P.J.;
 RA "The sister of P-glycoprotein represents the canalicular bile salt
 RT export pump of mammalian liver."
 RT J. Biol. Chem. 273:10046-10050(1998).
 RL J. Biol. Chem. 273:10046-10050(1998).
 CC -1- FUNCTION: INVOLVED IN THE ATP-DEPENDENT SECRETION OF BILE SALTS
 CC INTO THE CANALICULUS OF HEPATOCYTES.
 CC -1- SUBCELLULAR LOCATION: Integral membrane protein.
 CC -1- TISSUE SPECIFICITY: EXPRESSED PREDOMINANTLY, IF NOT EXCLUSIVELY IN
 CC THE LIVER, WHERE IT WAS FURTHER LOCALIZED TO THE CANALICULAR
 CC MICROVILLI AND TO SUBCANALICULAR VESICLES OF THE HEPATOCYTES BY IN
 CC SITU.
 CC -1- DOMAIN: MULTIFUNCTIONAL POLYPEPTIDE WITH TWO HOMOLOGOUS HALVES,
 CC EACH CONTAINING AN HYDROPHOBIC MEMBRANE-ANCHORING DOMAIN AND AN
 CC ATP BINDING CASSETTE (ABC) DOMAIN.
 CC -1- SIMILARITY: BELONGS TO THE ABC TRANSPORTER FAMILY. MDR SUBFAMILY.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL: U69487; AAC0084.1;
 CC InterPro: IPR003593; AAA_ATPase.
 CC InterPro: IPR003439; ABC_transport.
 CC InterPro: IPR001140; ABCtransprtm.
 CC Pfam: PF00005; ABC_tran; 2.
 CC Pfam: PF00664; ABC_membrane; 2.
 CC ProDom: PD000006; ABC_transport; 2.
 CC SMART: SM00382; AAA; 2.

DR PROSITE: PS00211; ABC_TRANSPORTER; 1.
 KW ATP-binding; Transmembrane; Transport.
 FT DOMAIN 1 62
 FT TRANSMEM 63 83
 FT DOMAIN 84 147
 FT TRANSMEM 148 168
 FT DOMAIN 169 215
 FT TRANSMEM 216 236
 FT DOMAIN 237 240
 FT TRANSMEM 241 261
 FT DOMAIN 262 319
 FT TRANSMEM 320 340
 FT DOMAIN 341 353
 FT TRANSMEM 354 374
 FT DOMAIN 375 755
 FT TRANSMEM 756 776
 FT DOMAIN 777 794
 FT TRANSMEM 795 815
 FT DOMAIN 816 869
 FT TRANSMEM 870 890
 FT TRANSMEM 891 911
 FT DOMAIN 912 979
 FT TRANSMEM 980 1000
 FT DOMAIN 1001 1011
 FT TRANSMEM 1012 1032
 FT DOMAIN 1033 1321
 FT NP_BIND 455 462
 FT NP_BIND 1113 1120
 FT CAROHD 109 109
 FT CAROHD 116 116
 FT CAROHD 122 122
 FT CAROHD 125 125
 SQ SEQUENCE 1321 AA; 146257 MW; 543F4E7B9F51F6 CRC64;
 Query Match 49.5%; Score 3205.5; DB 1; Length 1321;
 Best Local Similarity 50.3%; Pred. No. 1,76-173;
 Matches 647; Conservative 233; Mismatches 372; Indels 35; Gaps 11;
 QY 20 KSKRNEK-KKKRPVSTFAMFRYSNMILRLVMTLVTGAALIHCAALPLMLVFGNMVDS 78
 DB 30 KSRQDGMKMKDRIKRGVFFELFRSSSKKIMLMGVICALHGMQPGILIGIMTDI 89
 QY 79 FANAGISR-----KTEPVILNES-----ITNTQHFIMHLEBMTTAYYSGI 123
 DB 90 FTKYIDENQLEITFGKACVNNLT-VWINSFPHONTNCTGVLADISEMIRKSGIAGV 148
 QY 124 GAGVLVAAYIVGSFWSCLAAGRQILTKRQFHAIMRQELGWDVHDVDELNTRLDVSK 183
 DB 149 GMYVLILGTFQIRLMVITGARQIRRMKTYFRIRIMKEIGWFDCTSVGELNSRFADIEK 208
 QY 184 IIEGIGDKIGMFHSIATFFTGIVGFTGMYLTVIILAIISPVGLSAIMAKILSSFTD 243
 DB 209 INDATLDQHLRFQRMSTANCGLLGFRGKRLTVILAVSPPLIGICAAVIGISAKFTE 268
 QY 244 KELLVAAAGVVAEYVLAIRVYIAGCGKRELYENKNELEANGIGIKATIANSTIGA 303
 DB 269 LELKRYAAGSIADEVLSIRVYAAFGGKNEKVEETENLEFARQWIMGKMGCFITCY 328
 QY 304 AELLIVASALAEWYGTSLVLSSEYSIGVLYTFEVSYLGAFSIGQASPIEAFANANG 362
 DB 329 MWCLTFEYALAEWYGTSLVLEDEETPGTIVOLFVILAAIMNIGHASSCLEIFSTGCS 388
 QY 363 AAYELFKIIDKPSIDBSYSGHRRPNIRKGNLEFNHVSFSPSKREKIKGALNKVQSG 422
 DB 389 AATNIFQIRIDRQPIIDMGSGDGKIDRIKGELEFHNATYFHPSPDKILDNLSWYIKRG 448
 QY 423 QTVLVNNGSGGKSTYVQLMQLIDPITDQVNCIDGDIRITINVAHLREITGVYSGEPLVF 482
 DB 449 ETTALVNGSGGKSTALQILORFYDPCGEMWYTLGDHDIRSLNTIMKLRDQIGYDEEPLVF 508
 QY 483 ATTAAENIRYGRENVYMTDETERAWEANAAYDFIMKLPPKPEITVIGERGAQLSGGOKORIA 542
 DB 509 STTAAENIRFGREDATMEDIVQAADANAYNFIYALPPQFDTLVGEQGGMGSGGOKORIA 568

```

OY 543 IARALVNRPKILLDEATSAIDTESEAVVOVALDKARKGRTTIVIAHRLSTVENADVANG 602
DB 569 IARALIRNPKILLDMATISALDNSEARVOALNKHIGHITISVARIHSTVRAADVITIG 628
OY 603 FDDGVAVEKGNIDELMKRKGIVFKLVMTQGRNEIELENA-TGESSESDALDM-----S 656
DB 629 FEHGAVERGHEELERKGVFMVLVTLOSOGNAHETSIMKDKATREGTLEKRFSGS 688
OY 657 PDSSGSLIKRSTFR--SIHAP--QGDRKLGCTKD-----LNNVPVSFWRLTKN 706
DB 689 YDLSIRASIRORSQSLSLTHDPPLAVADHKSYSKDSKNDVIAEVEPAPVRRLTKN 748
OY 707 STEMYFVVGJFCALINGLOPAPSIIFSRIGITFREDEPETKRONSMFVLVIGI 766
DB 749 IPEMHYILVGSLSAINGAVPIYSLLFSQLGTFSL-LDKEQSEHSMCLFVVLIGC 807
OY 767 ISFTFFLOGTFPGAGHILTRKRYVFRSMKROVSWFDDPKNTGALTRLANDAAQ 826
DB 808 VSIFGLOGTTFKSGSELLKRLKRGFKAMGODIGFDDLRRNPGVLTTRLATASQ 867
OY 827 VKGAGSLAVTTQNIANLGTGIIISLIYGMOTLLTLLAIPILIAAGVEMKMLSGAL 886
DB 868 VOGATGSOVMMVMSFNITIALLLAFPSWKLITITFFPPLALSGAVOTKMLTGFAS 927
OY 887 KDKELGAGKINTEALENFRYVSLTRQKFEYMAQSLQVRYNSLRKAHIFGSFSI 946
DB 928 QDKQALFKAGQITSEALSNIRYVAGVGEGRIFKAFEVELQTSKTAVRKANITIGLCFAF 987
OY 947 TQAMMYFSYAGCFRGAVIVANEFMFOVLLVFSAIYFGAMAYGOVSSFPDYAKAVS 1006
DB 988 SSGIAFLANSAAVRYGCIILAYEGIGFSHVRVVSVALSTAAGRFSSTPSTAKAKIS 1047
OY 1007 AAHVIMIKSPILDSYSHGKPKMTLGCVNTFNEVFNYPTRPDIVLGLSLEVAKG 1066
DB 1048 AAFEPOLDRKPPILNYSIPAGEKMDNFGKIDFIDCKFTYPSRDPIDVNLGVSANPQ 1107
OY 1067 TIALVSSGCGKSTVOLLERFYPDLASVLDGKEIKHLNVOMLRHLGVQOEPILD 1126
DB 1108 TIAFVSSGCGKSTIGLERFYPDQGTVMIDGHDKVNIIDFLRSNIGIYVQOEVLDP 1167
OY 1127 CSTAENIYAGDSRYVSHHEEIMOKAKENIHFTLETPERYNTRVYGDGTGSGGOKRI 1186
DB 1168 CSTMONIKGDMTKETISVERAIAAKOALHDFWMSLPEKYEINVGIOGQSREGKOKI 1227
OY 1187 AIARALVROPHILLDEATSAIDTESKVVQVALDKAREGRTIYIAHRLSTQONADLV 1246
DB 1228 AIARALVROPHILLDEATSAIDTESKRTVQVALDKAREGRTIYIAHRLSTQONADLV 1287
OY 1247 VFQNGKVERHGTTHOOLAKGITYFSWV 1273
DB 1288 VVSGGVATIEKGTHERKMAOKGAYYKLV 1314

RESULT 15
MDRL_CAEEL STANDARD; FRT: 1321 AA.
ID MDRL_CAEEL STANDARD; FRT: 1321 AA.
AC P34712;
DB 01-FEB-1994 (Rel. 28, Created)
DB 01-FEB-1994 (Rel. 28, Last sequence update)
DB 16-OCT-2001 (Rel. 40, Last annotation update)
DB Multidrug resistance protein 1 (P-glycoprotein A).
DB Pgp-1.
DB Caenorhabditis elegans.
DB Eukaryota; Metazoa; Chromadorea; Rhabditida; Rhabditidae;
DB Rhabditidae; Pelodierinae; Caenorhabditis.
DB NCBI_TaxID=6239;
DB [1]
DB SEQUENCE FROM N.A.
DB STRAIN=Bristol N2;
DB MEDLINE=93085750; PubMed=1360540;
DB Lincke C.R., The I. van Groenigen M., Borst P.;
DB "The P-glycoprotein gene family of Caenorhabditis elegans. Cloning

```

```

RT and characterization of genomic and complementary DNA sequences.";
RT J. Mol. Biol. 228:701-711(1992).
RN [2]
RP TISSUE SPECIFICITY.
RC STRAIN=Bristol N2;
RA MEDLINE=93223702; PubMed=8096815;
RX Lincke C.R., Broeks A., The I., Plasterk H.A., Borst P.;
RT "The expression of two P-glycoprotein (pgp) genes in transgenic
RT Caenorhabditis elegans is confined to intestinal cells.";
RT EMBO J. 12:1615-1620(1993).
CC - FUNCTION: ENERGY-DEPENDENT EFFLUX PUMP RESPONSIBLE FOR DECREASED
CC - DRUG ACCUMULATION IN MULTIDRUG-RESISTANT CELLS.
CC - TISSUE SPECIFICITY: INTEGRAL membrane protein.
CC - SIMILARITY: BELONGS TO THE ABC TRANSPORTER FAMILY, MDR SUBFAMILY.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DB EMBL: X65054; CAA6190.1; -.
DB PIR: S27337; S27337.
DB InterPro: IPR003593; AAA_ATPase.
DB InterPro: IPR003493; ABC_transport.
DB InterPro: IPR001140; ABCtransportTM.
DB Pfam: PF000005; ABC_tran; 2.
DB Pfam: PF00664; ABC_membrane; 2.
DB ProDom: PD000006; ABC_transport; 2.
DB SMART: SM00382; AAA; 2.
DB PROSITE: PS00211; ABC_TRANSPORTER; 2.
DB APB-binding; Glycoprotein; Transmembrane; Repeat;
DB Multigene family.
FT DOMAIN 1
FT TRANSMEM 78 77 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 144 164 POTENTIAL.
FT TRANSMEM 213 233 POTENTIAL.
FT TRANSMEM 240 260 POTENTIAL.
FT TRANSMEM 321 341 POTENTIAL.
FT TRANSMEM 350 370 POTENTIAL.
FT DOMAIN 371 753 POTENTIAL.
FT TRANSMEM 754 774 POTENTIAL.
FT TRANSMEM 798 818 POTENTIAL.
FT TRANSMEM 874 894 POTENTIAL.
FT TRANSMEM 895 915 POTENTIAL.
FT TRANSMEM 978 998 POTENTIAL.
FT TRANSMEM 1017 1037 POTENTIAL.
FT DOMAIN 1038 1321 CYTOPLASMIC (POTENTIAL).
FT NP_BIND 451 458 ATP (POTENTIAL).
FT NP_BIND 1112 1119 ATP (POTENTIAL).
FT SEQUENCE 1321 AA; 145074 MW; 6130AAFB3B1A8FA9 CRC64;

Query Match 42.3%; Score 2739; DB 1; Length 1321;
Best Local Similarity 45.0%; Pred. No. 4, 1e-147;
Matches 584; Conservative 221; Mismatches 434; Indels 58; Gaps 12;

OY 21 KSKRKEKKEKPYSTRAMRYSNMDRLVMTGMAAIHGAJPLIMLVGNNTDSFA 80
DB 47 KITRDAKEEVNKRVSIPOLVRYTTLEKLLFGLVAVITGAGLWMSILOGKRSQAFI 106
OY 81 NAGISRNKTFPVILINISITNNTOHFI-----NLSEEMTYAYVYSGAGVLAAY 132
DB 107 NE-----QIVIN-----NNGSTFPLPTQANTTKTDFEDHVNVMVSYAAMTVGMMAAQ 154
OY 133 IOVSFWCLAAGROILKIRKOFFHAIRKQETGFWFDVHVGELNTRLDVSKINGIGDKI 192
DB 155 IIVCYLYLVAEQNNNRIRREFFVKSIIROEISWFDNHSGLATKFLDLERVKRGTDKI 214
OY 193 GMEFHSIAFFETFIIVGTRGKLTLYLTAISPVLGSAALMAKILSSFDKELAYAKA 252

```

Search completed: December 9, 2002, 17:17:04
job time : 297 secs

Dh	215	GNMFOYLSOPITGEIVAFTHSMOLLTVMLAVTPIQALGCFALAKSMSPFAIRELIRAKNA	274
Oy	253	GNAVEVLAIRFTVIAFGGCKELERYNKNLEBANGCIGIKRATITANISIGAFELLIVASY	312
Dh	275	GKAVEETISITVIVSLANGLEYELERSTAVEBAKRAVYLKGLGIFSGAMQASNITSP	334
Oy	313	ALAEMVGTSLVLSSEYSIGOVLYVFSVVLGASIGQASPSITAEAFNARKGAAYEIFIID	372
Dh	335	ALAEYIGVGVHDSINFGDMLTFFSSVMGSMALGIAGPOJLAVLCTMOGAASOIEVID	394
Oy	373	NKPEIDSYKSGHKRPONIKGNLEFNKNFYSPSRKEEYKILKGLNKKVOSQOTVALVNSG	432
Dh	395	RKPIIDSSKAGKMDKIKGDIYVENHFTVPSRPDVLGRMNLRVNAGOTVALVSSG	454
Oy	433	CGKSTVOMLRYPDPTDMCICODIRITNVNHIAREITGVOSOPVLEFATIENTIRY	492
Dh	455	CGKSTIISLLRYVDLAKIGRIDVDVDRINLEFKKNVAVSOPALFNQTIENISL	514
Oy	493	GRENVTUDELERKAVKEANAYDTFMLPKPEFTLVBERGEAOLSGGOKORIALARLVNPK	552
Dh	515	GREGTIREMAACKANANEKRTIKLPLNGVMTVGDGOTLGGGOKORIALARLVNPK	574
Oy	553	ILLDEATSAIDEESEAVNOVALDKARKGRTTYIARLSTVNRNAVYIGFDDGYIVBK	612
Dh	575	ILLDEATSAIDEESEGIYOALDKARKGRTTYIAHLSTIRNADILISCKNGOVLEV	634
Oy	613	NHDELMKKGIFYLFLVMTORNGEIR-----LENATGSKSESADLEMSFKOSGSL	664
Dh	635	DHRALMAOGLIYDLVLAQTFPTVADSAEKGFSRENSVAROTSEHGSLRSQASEMDIM	694
Oy	665	IK-RRSTRRSI-HAPQ-----GDRKLGRKEDLNANPPVSPFNMILINSTEMPY--	712
Dh	695	NRVBSITGISITNGVIDEKEERIGDALSRUKQELSEN-----NAOKTNLEFLLVNA	747
Oy	713	-----FVVGJFCALITNGLOPAFSIISFISIGITFRDEDETRKONSMNPSVLFVLCI	766
Dh	748	RPHALSLEFGMSTANITGFIYPIYSVFETSSFNWFA--GNPADELSQGHWMLFVLVA	805
Oy	767	ISFTYFPELOGFTFGAGEIILKRLIVNPFMSMLRODYSMPDDPKMTGALTRLANDAO	826
Dh	806	AGGICGFLMTFPMGLASESLRLDRNLKLFNNVLSQIHGFSDPOASKISITRLATDVN	865
Oy	827	VKAGISRLAVITOMIANIGITISILYSYMOFTLLLTAVTIIAIGVEMKMSQAL	886
Dh	866	LRLAIDFRSIVYITTVLSVWAGICLAFYGMQAMALLIATLPIVAFQYLGRGRFTGNV	925
Oy	887	KDKKELEGAGKATEIAENEFVTVSLTBEOKEFYEMVAOSLOVPYNSIRKRAHIFGSEFI	946
Dh	926	KSMSEFMDSKIAIEIEVNRVFOALAEDETFEYENFCEKEDIDPHKAEAFIOGLSYGC	985
Oy	947	TOAMNYSYAGCRRFCAVYANEEFNNFOD--VLLVEFATIFGMAVGOVSPADPYAKAK	1004
Dh	986	ASSVLYLILNCAIRMGJALITIDPTPQPMNLVAMVAIITISLTGFAITSFEPEYAKAT	1045
Oy	1005	VSAAHVIMITEKSPALDIDSYSPHGLKPNTLEGNVTFNEVYNNYTRRDIDVYLOGSLYEKK	1064
Dh	1046	FAGGILGMLRKISIDLSLAGEK-KLYKRVIEFKVNRVAYERPELTIELGLSFSVER	1104
Oy	1065	GOTLALVSSGCGKSTYVOUTLEFPDPLAGSLVLDKEIKHILNOMLRAHLIGVSOEPTL	1124
Dh	1105	GOTLALVSSGCGSTYVALLERFDITDLOGELFIDSGISIKUNDEHTRSOAIVASOEPUL	1164
Oy	1125	PCOSIAENIAVAGDSRVVSHETMOAKAEANIHFETLPEKYITRVGKGOTLSSGGO	1184
Dh	1165	PCOSIAENIITGIDPSSVTMAOVEEAFARLIANHINIALPGEFTVRGDTGOTLSSGGO	1224
Oy	1185	RIARALAVROPHILLDEATSAIDTESKVOVALDKARGRFCIYIARLSTIONADL	1244
Dh	1225	RIARALAVRNPKIILLDEATSAIDTESKVOVALDKARGRFCIYIARLNTVMNADC	1284
Oy	1245	IYVPONGKVKHEGTHOOLLAOKGITYSMVSYOAKAR	1281
Dh	1285	IAYVNSGILIEKGHTOOLMSKAGVYKTLQOKWOTTEK	1321

GenCore version 5.1.3
Copyright (c) 1993 - 2002 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: December 9, 2002, 16:20:53 ; Search time 1313 Seconds

(without alignments)
201.025 Million cell updates/sec

Title: US-09-672-725C-2

Perfect score: 6477
Sequence: 1 MDPEGGRKSAEKNFWMKMGK.....LAAQKGYFSWVSQAGAKR 1281

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 671580 segs, 206047115 residues

Total number of hits satisfying chosen parameters: 671580

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Database : Listing first 45 summaries

1: sp.archaea:*
2: sp.bacteria:*
3: sp.fungi:*
4: sp.human:*
5: sp.invertebrate:*
6: sp.mammal:*
7: sp.mhc:*
8: sp.organelle:*
9: sp.phage:*
10: sp.plant:*
11: sp.podent:*
12: sp.virus:*
13: sp.vertebrate:*
14: sp.unclassified:*
15: sp.virus:*
16: sp.bacteriaph:*
17: sp.archaeap:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	6420.5	99.1	1280	6	046605
2	5682.5	87.7	1285	6	002793
3	5629	86.9	1272	11	09JRK64
4	5607	86.6	1272	11	0924L6
5	5516.5	85.2	1163	6	09T5U2
6	5333	82.3	1275	11	060502
7	5286.5	81.6	1275	11	08R4Z7
8	4660.5	72.0	1288	13	093437
9	4430	68.4	1287	13	091586
10	3331	51.4	1348	13	091035
11	3315	51.2	1348	13	091035
12	3209.5	49.6	1321	11	088331
13	2857	44.1	1294	5	045721
14	2852.5	44.0	1407	5	08T9W5
15	2827	43.6	1265	5	001495
16	2825	43.6	851	13	09W693

17	2818.5	43.5	1275	5	061301
18	2751.5	42.5	1432	5	08P9W4
19	2741	42.3	1321	5	021349
20	2701	41.7	1313	5	09Y6Z6
21	2652	40.9	1283	5	024393
22	2612.5	40.3	817	13	090WV6
23	2603.5	40.2	1285	10	09L6X1
24	2602.5	40.2	1286	10	080725
25	2596	40.1	1292	10	09M109
26	2589.5	40.0	1278	10	09FWX7
27	2557	39.5	1279	5	09V616
28	2540.5	39.2	1262	10	08RV7
29	2532	39.1	1348	3	09Y8G1
30	2521	38.9	1229	10	09SVY2
31	2518	38.9	1229	10	049749
32	2518	38.9	1230	10	09SVY3
33	2517	38.9	1289	10	094TH6
34	2511	38.8	1349	3	043121
35	2489.5	38.4	1287	10	08S035
36	2473.5	38.2	1313	10	09ZRG2
37	2473	38.2	1229	10	09FWX8
38	2465	38.1	1292	3	09C163
39	2441	37.7	1408	3	043140
40	2437.5	37.6	1323	10	09W0M2
41	2431	37.5	1286	10	09ZKR7
42	2423	37.4	1310	5	024852
43	2419.5	37.4	1248	10	09FHF1
44	2418.5	37.3	1331	3	09HGF5
45	2410	37.2	1252	10	09LX0

ALIGNMENTS

RESULT 1

046605 PRELIMINARY; PRT: 1280 AA.
ID 046605:
AC 01-JUN-1998 (TREMBLER, 06, Created)
DT 01-JUN-1998 (TREMBLER, 06, Last sequence update)
DT 01-JUN-2002 (TREMBLER, 21, Last annotation update)
DE Multidrug resistance P-glycoprotein.
GN MDRL.
OS Canis familiaris (Dog).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Carnivora; Fissipedia; Canidae; Canis.
OX NCBI_TaxID=9615;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=COLON;
RA Puel O., LePage J.F., Alvine M., Gallier P., Pineau T.;
RT Submitted (JAN-1998) to the EMBL/GenBank/DBJ databases.
CC -1- SIMILARITY: BELONGS TO THE ABC TRANSPORTER FAMILY.
DR EMBL: AF045016; AAC02113.1; -;
DR InterPro: IPR003593; AAA_Arpase.
DR InterPro: IPR001140; ABCtransporter.
DR InterPro: IPR003439; ABCtransporter.
DR Pfam: PF00664; ABC_membrane_2.
DR Pfam: PF00005; ABC_tran; 2.
DR ProDom: PD000006; ABC_transport; 2.
DR SMART: SM00382; AAA; 2.
DR PROSITE: PS00211; ABC_TRANSPORTER; 2.
DR ATP-binding; Transport.
KW SEQUENCE 1280 AA; 141524 MW; 762DD5AF4C73306 CRC64;

Query Match 99.1%; Score 6420.5; DB 6; Length 1280;
Best Local Similarity 99.3%; Pred. No. 0;

Matches 12/2; Conservative 4; Mismatches 4; Indels 1; Gaps 1;

OY 1 MDPEGGRKSAEKNFWMKMGKSKKNEKKKPTVSTFAMFRYSNWLDRLTYLVGTMAII 60
Db 1 MDPEGGRKSAEKNFWMKMGKSKK-EKKKKPTVSTFAMFRYSNWLDRLTYLVGTMAII 59

61 HGALPLMLVFGNMTDSFANAGISRNKTPVITNESITNNQHFINHLEEMTYAYXX 120
DB 60 HGALPLMLVFGNMTDSFANAGISRNKTPVITNESITNNQHFINHLEEMTYAYXX 119
QY 121 SGIGAGVVAAYIOVSFMCAGROILKIRKOFPHAIMROEIGWFDVHDVDELNRLD 180
DB 120 SGIGAGVVAAYIOVSFMCAGROILKIRKOFPHAIMROEIGWFDVHDVDELNRLD 179
QY 181 VSKINEGIGDKIGMEFHSIATFPGTIVGTGKGLVLVLAISPVIGLSAAILAKILSS 240
DB 180 VSKINEGIGDKIGMEFHSIATFPGTIVGTGKGLVLVLAISPVIGLSAAILAKILSS 239
QY 241 FPDKRELLAARAGAAVEVLAIRTVIAFGOKKELERYNKNLEAKGIGIKKAITANIS 300
DB 240 FPDKRELLAARAGAAVEVLAIRTVIAFGOKKELERYNKNLEAKGIGIKKAITANIS 299
QY 301 IGAAPLLIYASALAFWYGTSLVSESYSTQVLTVPFVSUIGAFSIGQASPSIEAFANA 360
DB 300 IGAAPLLIYASALAFWYGTSLVSESYSTQVLTVPFVSUIGAFSIGQASPSIEAFANA 359
QY 361 RGAAYIEFKIIDNKPSIDSYSGSKHPDNKGNLEFKVHFSYSPREKVKILKGLNLKVQ 420
DB 360 RGAAYIEFKIIDNKPSIDSYSGSKHPDNKGNLEFKVHFSYSPREKVKILKGLNLKVQ 419
QY 421 SGGTVALVNSGCGKSTVQLMQRLYDPTGAVCIDGODIRTYNRLREITGYVSOEPV 480
DB 420 SGGTVALVNSGCGKSTVQLMQRLYDPTGAVCIDGODIRTYNRLREITGYVSOEPV 479
QY 481 LEATTIAENIRIGRENVYDELEKAVKEMANADFTMKLPNKFDTLVGERGARLSGGQKOR 540
DB 480 LEATTIAENIRIGRENVYDELEKAVKEMANADFTMKLPNKFDTLVGERGARLSGGQKOR 539
QY 541 IATARALVNRPKILLDEFTSALDTESEAVOVALDKARKGRTTIVIAHRLSTVNRADVI 600
DB 540 IATARALVNRPKILLDEFTSALDTESEAVOVALDKARKGRTTIVIAHRLSTVNRADVI 599
QY 601 AGEDDGVIVEKGNHDELMKEKGYEFVLVMQTRNGEIELENAITGSKESDALESPDS 660
DB 600 AGEDDGVIVEKGNHDELMKEKGYEFVLVMQTRNGEIELENAITGSKESDALESPDS 659
QY 661 GSSILIKRSTRRSIHAPOGODRKLGTKEDELENVPPVSFWRLILKINSTEMPYFVVGIFCA 720
DB 660 GSSILIKRSTRRSIHAPOGODRKLGTKEDELENVPPVSFWRLILKINSTEMPYFVVGIFCA 719
QY 721 IINGLOPAFSIIFSRILIGITRDEDEPFRKONSMEVFLVYLGLISTITFLOGFTFG 780
DB 720 IINGLOPAFSIIFSRILIGITRDEDEPFRKONSMEVFLVYLGLISTITFLOGFTFG 779
QY 781 KAGELIKRLRYMYERSMRLRODYSWPDPKNTGALTTRLANDAQAIGSRLAVITQ 840
DB 780 KAGELIKRLRYMYERSMRLRODYSWPDPKNTGALTTRLANDAQAIGSRLAVITQ 839
QY 841 NIANGIGIILISLIGMQLTLLLAIVPITAIAGVEMKMLSGOALKDKKELEGAKIAT 900
DB 840 NIANGIGIILISLIGMQLTLLLAIVPITAIAGVEMKMLSGOALKDKKELEGAKIAT 899
QY 901 EAIENFRVVSILTRBOKFEYMYASLOVPRNSLKAHITGVSFSTIOAMMYSTVAGCRR 960
DB 900 EAIENFRVVSILTRBOKFEYMYASLOVPRNSLKAHITGVSFSTIOAMMYSTVAGCRR 959
QY 961 FGAYIVANEFNFDVLIVFSAIFGAMAVGOVSSFAADYAKAKVASAHYIMITKESPLI 1020
DB 960 FGAYIVANEFNFDVLIVFSAIFGAMAVGOVSSFAADYAKAKVASAHYIMITKESPLI 1019
QY 1021 DSYSPHGLKPTLBNVTTFNEVYNYPTRPDIPLVLOGLSIEVKKGGOTLAVGSSGCKST 1080
DB 1020 DSYSPHGLKPTLBNVTTFNEVYNYPTRPDIPLVLOGLSIEVKKGGOTLAVGSSGCKST 1079
QY 1081 VVOLLERFYDPLAGSVLIDGKEIKHLNOMLRAHGIIVSOEPIIFDCSIAENIAYGDNRR 1140
DB 1080 VVOLLERFYDPLAGSVLIDGKEIKHLNOMLRAHGIIVSOEPIIFDCSIAENIAYGDNRR 1139
QY 1141 VVSHEEIMQAKKANIHHFIETLPEKTYTRVGDGKTOLSGGQKORIALAIALVROPHILL 1200

DB 1140 VVSHEEIMQAKKANIHHFIETLPEKTYTRVGDGKTOLSGGQKORIALAIALVROPHILL 1199
QY 1201 IDENTSALDTESEKVOALDKAREGRTCIIVAHRLSTIONADLIYVPONGVKEHGTHQ 1260
DB 1200 IDENTSALDTESEKVOALDKAREGRTCIIVAHRLSTIONADLIYVPONGVKEHGTHQ 1259
QY 1261 QLLAOKGIFYSMVSQAGAKR 1281
DB 1260 QLLAOKGIFYSMVSQAGAKR 1280
RESULT 2
002793 PRELIMINARY; PRT; 1285 AA.
ID 002793
AC 002793
DT 01-JUL-1997 (TREMBLrel. 04, Created)
DI 01-JUL-1997 (TREMBLrel. 04, Last sequence update)
DT 01-JUN-2002 (TREMBLrel. 21, Last annotation update)
DE Multidrug resistance protein-1.
GN MDRL.
OS Ovis aries (Sheep).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
OC Bovidae; Caprinae; Ovis.
OX NCBI_TaxId=9940;
RN [1]
RP SEQUENCE FROM N.A.
RA Longley M., Crawford A.M.
RT "Ovine mdrl gene."
RL Submitted (NOV-1996) to the EMBL/GenBank/DBJ databases.
CC -1- SIMILARITY: BELONGS TO THE ABC TRANSPORTER FAMILY.
DR EMBL: U78609; AAB58489.1; -
DR InterPro: IPR003593; AAA_Arase.
DR InterPro: IPR001140; ABCtransport.
DR InterPro: IPR003439; ABC_transport.
DR Pfam: PF00664; ABC_membrane_2.
DR Pfam: PD000006; ABC_tran; 2.
DR Prodom: PD000006; ABC_transport; 2.
DR SMART: SM00382; AAA; 2.
DR PROSITE: PS00211; ABC_TRANSPORTER; 2.
KW ATP-binding; Transport.
SQ SEQUENCE 1285 AA; 142020 MW; 90153617C44856F CRC64;
Query Match 87.7%; Score 5682.5; DB 6; Length 1285;
Best Local Similarity 87.3%; Pred. No. 0;
Matches 1122; Conservative 80; Mismatches 78; Indels 5; Gaps 3;
QY 1 MDPGGRKG--SARKNFMKMGKK---SKNNEKKKRPVSTFAMPRYSNMLDRLVNLVGT 56
DB 1 MDEGDRNGRRAGGNFLKRDKKRFFSKKDEKKEKRTVSTFTMFPRYSNMLDRLCYVLTGL 60
QY 57 AAIHGAALPLMLVFGNMTDSFANAGISRNKTPVITNESITNNQHFINHLEEMTY 116
DB 61 AAIHGAALPLMLVFGDMTDSFAGAGNLGNTILSNISTSTIDRTE-VGKLEKEMTY 119
QY 117 AAYTSGAGVVAAYIOVSFMCAGROILKIRKOFPHAIMROEIGWFDVHDVDELNRL 176
DB 120 AAYTSGAGVVAAYIOVSFMCAGROVYHRLKOFPHAIMROEIGWFDVHDVDELNRL 179
QY 177 LTDDVSKINEGIGDKIGMEFHSIATFPGTIVGTGKGLVLVLAISPVIGLSAAILAK 236
DB 180 LTNDVSKINEGIGDKIGMEFHSIATFPGTIVGTGKGLVLVLAISPVIGLSAAILAK 239
QY 237 ILSFTDKELLAYAKAGAAVEVLAIRTVIAFGOKKELERYNKNLEAKGIGIKKAIT 296
DB 240 ILSFTDKELLAYAKAGAAVEVLAIRTVIAFGOKKELERYNKNLEAKKRRIGIKKAIT 299
QY 297 ANISIGAAPLLIYASALAFWYGTSLVSESYSTQVLTVPFVSUIGAFSIGQASPSIEA 356
DB 300 ANISIGAAPLLIYASALAFWYGTSLVSESYSTQVLTVPFVSUIGAFSIGQASPSIEA 359
QY 357 FANARGAAYIEFKIIDNKPSIDSYSGSKHPDNKGNLEFKVHFSYSPREKVKILKGLN 416

Dd	360	FANAGAAVEYFKLINDKPSIDSYNTGHRPNIKGNLEFRVHEHYPSREVKILKGLN	419
Qy	417	LKVGSGGTVALVNGSGCGKSTTVQLMQRIYDPTDQVAGCIDGODIRTNVRLREITGVYS	476
Dd	420	LKVGSGGTVALVNGSGCGKSTTVQLMQRIYDPTDQVAGCIDGODIRTNVRLREITGVYS	479
Qy	477	QEPLEFPTTAENIRGRENVTWDEIEKAVKEANAIDEIMLIPKPFITVIGERQAQLSGG	536
Dd	480	QEPLEFPTTAENIRGREDVYTWDEIQKAVKEANAIDEIMLIPKPFITVIGERQAQLSGG	539
Qy	537	QKQRIATARALVNPRIKILLDEATSLDTESEAVVQVALDKARKGRFTTYIAHRLSTVRN	596
Dd	540	QKQRIATARALVNPRIKILLDEATSLDTESEAVVQVALDKARKGRFTTYIAHRLSTVRN	599
Qy	597	ADVIAGDDGVIVEKNHDELMKEKGIYFKYLVYMQTRGNEITELENATGEKSSQDLNLS	656
Dd	600	ADVIAGDDGVIVEEGSHDELMKGRGIFYKLVYMQTRGNEITELENATGEKSSQDLNLS	659
Qy	657	PKDGSLLIRKSTRRSHAPGOCORIKGTIEDLNENVPVSEWRILKLNSTEMPVYVG	716
Dd	660	SQDRSSLIRKSTRRSHAPGOCORIKGTIEDLNENVPVSEWRILKLNSTEMPVYVG	719
Qy	717	IFCAIINGLOPAFSIIFSRIGIITPRDEDEETKRONSNFSLFVYLGISTITFFLQ	776
Dd	720	IFCAIINGLOPAFSIIFSRIGIITPRDNDDETRKQNSNLFSLFVYLGISTITFFLQ	779
Qy	777	FTEGKAGELLTKRLNMFPSKMLRODVSWFDPKNTTGALTTRILANDAAQVGAIGSRLA	836
Dd	780	FTEGKAGELLTKRLNMFPSKMLRODVSWFDPKNTTGALTTRILANDAAQVGAIGSRLA	839
Qy	837	VITQNIANIGTIIISLIYQMQLTLLIATVPIIAINGVEMAMLSGQALKKCKELEAG	896
Dd	840	VITQNIANIGTIIISLIYQMQLTLLIATVPIIAINGVEMAMLSGQALKKCKELEAG	899
Qy	897	KATKAIENFTVVSILTREQFEYMAQSLQVPRNSLRAKIHFGVSFIITQAMMYESEA	956
Dd	900	KATKAIENFTVVSILTREQFEYMAQSLQVPRNSLRAHFGITSIITQAMMYESEA	959
Qy	957	GCFRRFAYLVAAEFNPDVLLVSALVFGAMAVGVSSFPADYAKAKVSAAHVIMITEK	1016
Dd	960	GCFRRFAYLVAAEFNPDVLLVSALVFGAMAVGVSSFPADYAKAKVSAAHVIMITEK	1019
Qy	1017	SPILDSISPHGLKPTLEGNATFEVFNPTRPDIPLVQGLSLWKKGOTLALVSGGC	1076
Dd	1020	IPILDSISPHGLKPTLEGNATFEVFNPTRPDIPLVQGLSLWKKGOTLALVSGGC	1079
Qy	1077	GKSTVOYLLERTPYPLAGSVLIDGKEIKHLNQMWLRAHLGIYSQEPILFDCSIAMENIAYG	1136
Dd	1080	GKSTVOYLLERTPYPLAGSVLIDGKEIKHLNQMWLRAHGIYSQEPILFDCSIAMENIAYG	1139
Qy	1137	DNSRVVSHHEEIMOAKEANIHHEFLEPLPEKYNTRNGDKTKQLSGQKORIAIARALVROP	1196
Dd	1140	DNSRVVSHHEEIMOAKEANIHHEFLEPLPEKYNTRNGDKTKQLSGQKORIAIARALVROP	1199
Qy	1197	HILLDEATSLDTESEKVOEALDKARGRTCIYIARHSLSTIONADLIYFONGKIKHEH	1256
Dd	1200	HILLDEATSLDTESEKVOEALDKARGRTCIYIARHSLSTIONADLIYFONGKIKHEH	1259
Qy	1257	GTHOOLLAQGIYFMSVQAGAKR 1261	
Dd	1260	GTHOOLLAQGIYFMSVQAGAKR 1264	
RESULT 3			
O9JK64			
ID	O9JK64	PRELIMINARY:	PRT, 1272 AA.
AC	O9JK64		
DT	01-OCT-2000 (Tremblrel. 15, Created)		
DT	01-OCT-2000 (Tremblrel. 15, Last sequence update)		
DT	01-JUN-2002 (Tremblrel. 21, Last annotation update)		
DE	Multidrug resistance protein 1a.		
EN	PGY1.		

OS *Rattus norvegicus* (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus
ON NCBI_TaxID=10116;
RX [1]
RP SEQUENCE FROM N.A.
RC STRAIN=MISTAR; TISSUE=LIVER;
RA Hovsveld G.J.E.J., Wilms J.W.J., Hagenbuch B., Jansen P.L.M.,
RA Meijer D.K.F., Muller M.;
RT "Cloning and functional characterization of the rat multidrug
RT resistance protein Mdr1a."
RL Submitted (Apr-2000) to the EMBL/GenBank/DBJ databases.
CC -1- SIMILARITY: BELONGS TO THE ABC TRANSPORTER FAMILY.
DR EMBL, AF257746; AAF69007.1; -
DR InterPro: IPR003593; AAA_ATPase.
DR InterPro: IPR001140; ABCtransp1TM.
DR InterPro: IPR003439; ABC_transportr.
DR Pfam: PF00664; ABC_membrane; 2.
DR Pfam: PF00005; ABC_tran; 2.
DR Prodom: PD00006; ABC_transportr; 2.
DR SMART: SM00382; AAA; 2.
DR PROSITE, PS00211; ABC_TRANSPORTER; 2.
DR ATP-binding; Transport.
KW ATP-binding; Transport.
SQ SEQUENCE 1272 AA; 140326 MW; B5C1DE89FC986C5 CRC64.

Query Match	86.6%;	Score 5607;	DB 11;	Length 1272;
Best Local Similarity	86.1%;	Pred. No. 9, 2e-317;		
Matches 1103;	Conservative 80;	Mismatches 88;	Indels 10;	Caps
QY	1	MDPEGKRSASAKNNKMKGKKSKNEKKEKPPVSEFAFRKSNMLDRLYMLVGTMAAI	60	
Db	1	MELEEDLNRRARKNSKMGKSKK-EKKEKKPAVSYLTRFRAGMLDRPYMLGTIAAI	59	
QY	61	HGAAITPLMLVFGNMTDFSPANAGISRNKTEPVIIINESITNNQHPNHLSEEMTYAAY	120	
Db	60	HGIAIPLMLVFGDMTDFSPANGNRRSTSF-----YNATDIYARLEDEMTYAIY	110	
QY	121	SGIGAGVLYAAITQVSPFCNLAGROILKIRKOPFHAIIMPOEIQMFYHVGELNTRLDD	180	
Db	111	TGIGAGVLYAAITQVSLWCLAGRIHRIKQKPFHAIIMNOICMFPYHVGELNTRLDD	170	
QY	181	VSKINIEGDKIMEFHSIAFFETFGFIVGTTGRKLTPLVILAIIPVLGISAIAIKIIS	240	
Db	171	VSKINIEGDKICMFPQAAATFEFGFIIIGTFGRKLTPLVILAIIPVLGISAGIMAKIIS	230	
QY	241	FTDKELLAYKACAAVEYLAIIRPVIAFGGOKKELERYNKNLEEKGGIKKAITANIS	300	
Db	231	FTDELOAYKAAVAEYLAIIRPVIAFGQKKELERYNNNEPAKRIGIKKAITANIS	290	
QY	301	IGAAFLITVASYALAFWYGTSLVLSSEXSIGQVLTVEFSYLIGAFSTGASPSIEAFANA	360	
Db	291	MGAFLITVASYALAFWYGSILVISEKTIQGVLTVEFSYLIGAFSGQASPMIEAFANA	350	
QY	361	RGAAVEIFKIIDNKPISDYSKSGHKPDKINICNLEFKVHFSVSPREKVIILKGLLKY	420	
Db	351	RGAAVEFHSIIDNKPISDYSKSGHKPDKIQCNLEFKKIHFSYPSRKVDYILKGLLKYK	410	
QY	421	SGQVYALVNGSGCKSTTVYLMQRLTDPDQGWVCIDGODIPTINVRHLREITGVSOEY	480	
Db	411	SGQVYALVNGSGCKSTTVYLLQRLTDPJEGHVSIDGODIPTINVRHLREITGVSOEY	470	
QY	481	LFATTIAENIRYGRENTYMDIEIKAYEKANAIDYIMKLPNKEPDLVBERGAGOLSGGOKOR	540	
Db	471	LFATTIAENIRYGRENTYMDIEIKAYEKANAADYIMKLPNKEPDLVBERGAGOLSGGOKOR	530	
QY	541	IAIAAIVRNKTIILDEATSALDTSESEAVVOALDKARKGRTTVIAIHRISTVRADVI	600	
Db	531	IAIAAIVRNKTIILDEATSALDTSESEAVVOALDKARKGRTTVIAIHRISTVRADVI	590	
QY	601	AGPDDGVIVKGNHDELMKEKGIYFKLVMTQTRGNIELENAITGSEKSESSEDALEMSKOS	660	
Db	591	AGPDDGVIVKGNHDELMKEKGIYFKLVMAQTAAGNIELEGNACSKSGCIGNDMSSKOS	650	
QY	661	GSSLIKRSTRSRSHIAQGGDRKIGTRKEDLENVNPVPSFWRLTKINSTWEPYFVYGICA	720	
Db	651	GSSLIKRSTRSRSHIAQGGDRKIGTRKEDLENVNPVPSFWRLTKINSTWEPYFVYGICA	710	
QY	721	IINGGLQPAFSTISRTIGITPDRDEPFRKUNSNKSVLVLVGLISFTITFFIQGTFFG	780	
Db	711	IINGGLQPAFSTISRTKVVGVPTKNDTDEIQRONSINLSLFTLIGISFTITFFIQGTFFG	770	
QY	781	KAGEIITKRLKRWRSMLRODVSMEDPDKNTGALTTRLANDAQAQVGAIGSRILAVITQ	840	
Db	771	KAGEIITKRLKRWRSMLRODISMEDPDKNTGALTTRLANDAQAQVGAIGSRILAVITQ	830	
QY	841	NINMLGTGIIISLITGMOLTTLLLAIVPIAIAGVYEMKMLSGQALKDKKLEEGSKATAT	900	
Db	831	NINMLGTGIIISLITGMOLTTLLLAIVPIAIAGVYEMKMLSGQALKDKKLEEGSKATAT	890	
QY	901	EAIENFRVAVSLTRQKFEPIYAQSLQVPPYRNSLIRKAHITFGVSFSTQMMWFSVAGCFR	960	
Db	891	EAIENFRVAVSLTRQKFEPIYAQSLQVPPYRNSLIRKAHITFGVSFSTQMMWFSVAGCFR	950	
QY	961	FGAIVYANENMNPQDILVESAIVGAAGAAVGVSSFPADKAKKASAAHVIIMIEKSPLI	1020	
Db	951	FGAIVYANENMNPQDILVESAIVGAAGAAVGVSSFPADKAKKASASHIIRIEKPEIPI	1010	
QY	1021	DATSPHILKNTLEGNVTFNEVVFVNPTRPDIPIVGLGSLFKKQGTALVYSSSGCGKST	1080	

```
Db 1011 DSYTBGKJRMLEGNKFNKGVMEFPTRENPVLOGLSLEVKGGOTLALVSSCGCKST 1070
QY 1081 VVOLLERFYDPLAGSVLIDGKEIKHLNVOMLRAHLGIVSOEPLTDCSIAENIAGDMSR 1140
Db 1071 VVOLLERFYDPMAGTVFLDCKEIKOLNVOMLRAHLGIVSOEPLTDCSIAENIAGDMSR 1130
QY 1141 VVSHHEELMOAKKAAKANHHPETLPEKYNTRVGDKGTOLSGGOKORAIARALVROPHTLL 1200
Db 1131 VVSHHEELVRAKAAKANHHPETLPEKYNTRVGDKGTOLSGGOKORAIARALVROPHTLL 1190
QY 1201 LDEATSAIDTESKVVQVQALDRAKREGRCTIVAHRLSTIQANDLIVFPQNGKVEHGTQ 1260
Db 1191 LDEATSAIDTESKVVQVQALDRAKREGRCTIVAHRLSTIQANDLIVFPQNGKVEHGTQ 1250
QY 1261 QLLAQKGIYFSMVSVOAGAKR 1281
Db 1251 PLLAQKGIYFSMVSVOAGAKR 1271
```

RESULT 5

```
Q9TSU2 PRELIMINARY: PRT: 1163 AA.
AC Q9TSU2;
DT 01-MAY-2000 (Tremblrel. 13, Created)
DT 01-MAY-2000 (Tremblrel. 13, Last sequence update)
DT 01-JUN-2002 (Tremblrel. 21, Last annotation update)
DE Multi-drug resistance related (Fragment).
OS Felis silvestris catus (Cat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Carnivora; Fissipedia; Felidae; Felis.
OX NCBI_TaxID=9685;
RN [1]
RP SEQUENCE FROM N.A.
RA Okada Y., Nakamura N., Matsusiro H., Kato H., Setoguchi A., Yazawa M.,
RA Okada M., Wataai T., Hasegawa A., Tsujimoto H.;
RT "Molecular analysis of multidrug resistance in lymphoma cells in the
RT cat.";
RL Am. J. Vet. Res. 0-0-0(1999).
CC -1- SIMILARITY: BELONGS TO THE ABC TRANSPORTER FAMILY.
DR EMBL; AB029153; BA87071.1; -.
DR InterPro; IPR003593; AAA_ATPase.
DR InterPro; IPR001140; ABCtransportTM.
DR Pfam; PF00664; ABC_membrane_2.
DR Pfam; PF00005; ABC_tran_2.
DR ProDom; PD000006; ABC_transport_2.
DR SMART; SM00382; AAA_1.
DR PROSITE; PS00211; ABC_TRANSPORTER; 2.
KW ATP-binding; Transport.
FT NON_TER 1
FT TER 1163
SQ SEQUENCE 1163 AA; 128511 MW; 1B5B413776A93A26 CRC64;
```

Query Match 85.2%; Score 5516.5; DB 6; Length 1163;
Best Local Similarity 94.2%; Pred. No. 1.4e-311;
Matches 1096; Conservative 34; Mismatches 33; Indels 1; Gaps 1;

```
QY 61 HGALPLPLMLVFGNMTDSFNAGISRNKTFPYLINESITNNQHPINHLSEEMTYAYYY 120
Db 1 HGALPLPLMLVFGNMTDSFNAGISRNKTLTTLNITGESIYNDG-YEINRLEEMTYAYYY 59
QY 121 SGIGAGVLAAYIYVSWFCLAGRQILKIRKQFHAIMROEIGMFDVADVGLNRLDD 180
Db 60 SGIGAGVLAAYIYVSWFCLAGRQILKIRKQFHAIMROEIGMFDVADVGLNRLDD 119
QY 181 VSKINGIGDKTGMFPHSHATFTFTGIVFTGKMLTVLILASIPVLGLSAIWAAILSS 240
Db 120 VSKINGIGDKTGMFPHSHATFTFTGIVFTGKMLTVLILASIPVLGLSAIWAAILSS 179
QY 241 FTDKELLAYAKGAAVEEYLAIRIVYAFGGOKLEBYNNLEBAKIGIKKAITANIS 300
Db 180 FTDKELLAYAKGAAVEEYLAIRIVYAFGGOKLEBYNNLEBAKIGIKKAITANIS 239
```

```
QY 301 IGAFLITVASYALAWYGTSLVLSSEYSIGVLTWFFSVLIGAFSGOASPIAPANA 360
Db 240 IGAFLITVASYALAWYGTSLVLSSEYSIGVLTWFFSVLIGAFSGOASPIAPANA 299
QY 361 RGAAYEFKTIIDKPSIDSYSGKHPDIKGNLEKKNVHESYPSRKEVKILKGLNKVQ 420
Db 300 RGAAYEFKTIIDKPSIDSYSGKHPDIKGNLEKKNVHESYPSRKEVKILKGLNKVQ 359
QY 421 SGQVALVNSGGCGSTVQOLMORLYDPDGMVCIDGODIRITINVRHLREITGVSOEPP 480
Db 360 SGQVALVNSGGCGSTVQOLMORLYDPDGMVSIIDGODIRITINVRHLREITGVSOEPP 419
QY 481 LEFTTIAENIRYGRENVYMDIEKAAYKANAVDTMKLPNKEDLVGRGAOLSGGOKOR 540
Db 420 LEFTTIAENIRYGRENVYMDIEKAAYKANAVDTMKLPNKEDLVGRGAOLSGGOKOR 479
QY 541 IATARALVNPKITLLDEATSAIDTESKVVQVQALDRAKRGRTTIVAHRLSTIRNADVI 600
Db 480 IATARALVNPKITLLDEATSAIDTESKVVQVQALDRAKRGRTTIVAHRLSTIRNADVI 539
QY 601 AGFDGVIYKGNHDELMKEKGIYKLVYMTQNRNEIELENAAGESKESDPALEMSPKDS 660
Db 540 AGFDGVIYKGNHDELMKEKGIYKLVYMTQNRNEIELENAAGESISIDALEMSPKDS 599
QY 661 GSSLKRSTRSIHAPGODRKIGTKEDLMENVPVPSFWRLKLNSTEMPYFVVGIFCA 720
Db 600 GSSLKRSTRSIHASOGODRKIGTKEDLMENVPVPSFWRLKLNSTEMPYFVVGIFCA 659
QY 721 IINGLOPAFSTIFSRIGITFRDEDPETKRONSMFSVLEIVGITSITFTFLOGFTTG 780
Db 660 IINGLOPAFSTIFSRIGITFRDEDPETKRONSMFSVLEIVGITSITFTFLOGFTTG 719
QY 781 KAGELITRLRYNFRSMLRODVSFDDPKNTGALTTRLANDAQAQKAGISRLAVITQ 840
Db 720 KAGELITRLRYNFRSMLRODVSFDDPKNTGALTTRLANDAQAQKAGISRLAVITQ 779
QY 841 NIANGITIIISLYGQOLITLLAIVPITAIAGVEMKMLSGQALDKKELGACKIAT 900
Db 780 NIANGITIIISLYGQOLITLLAIVPITAIAGVEMKMLSGQALDKKELGACKIAT 839
QY 901 EAIENFTVSVLREOKFEYVAQSLQVPRNSLRKAHLGIVSFTSQAMMPFSYAGCGR 960
Db 840 EAIENFTVSVLREOKFEYVAQSLQVPRNSLRKAHLGIVSFTSQAMMPFSYAGCGR 899
QY 961 FGATLVANPEPNRQDVLVLSAIVFGMAVGOVSSFPADYAKAKVSAHVIMITEKSPLI 1020
Db 900 FGATLVANPEPNRQDVLVLSAIVFGMAVGOVSSFPADYAKAKVSAHVIMITEKSPLI 959
QY 1021 DSYSPHGLKPNTEGNTFENEVEFNTPTRPDIPLYLOGLSLEVKGGOTLALVSSCGCKST 1080
Db 960 DSYSPHGLKPNTEGNTFENEVEFNTPTRPDIPLYLOGLSLEVKGGOTLALVSSCGCKST 1019
QY 1081 VVOLLERFYDPLAGSVLIDGKEIKHLNVOMLRAHLGIVSOEPLTDCSIAENIAGDMSR 1140
Db 1020 VVOLLERFYDPMAGTVFLDCKEIKHLNVOMLRAHMGIVSOEPLTDCSIAENIAGDMSR 1079
QY 1141 VVSHHEELMOAKKAAKANHHPETLPEKYNTRVGDKGTOLSGGOKORAIARALVROPHTLL 1200
Db 1080 VVSHHEELVRAKAAKANHHPETLPEKYNTRVGDKGTOLSGGOKORAIARALVROPHTLL 1139
QY 1201 LDEATSAIDTESKVVQVQALDRAK 1224
Db 1140 LDEATSAIDTESKVVQVQALDRAK 1163
```

RESULT 6
Q60502 PRELIMINARY: PRT: 1169 AA.
AC Q60502;
DT 01-NOV-1996 (Tremblrel. 01, Created)
DT 01-NOV-1996 (Tremblrel. 01, Last sequence update)
DT 01-JUN-2002 (Tremblrel. 21, Last annotation update)

DE P-glycoprotein.
GN Pgp-1.
OS Cricetus sp.
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Cricetinae;
OC Cricetus.
OC NCBI_TaxID=10031;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=LUNG;
RX MEDLINE=91154265; PubMed=1671863;
RA Devine S.E., Hussain A., Davide J.P., Melera P.W.;
RT "Full length and alternatively spliced pgp-1 transcripts in multidrug
resistant Chinese hamster lung cells."
RL J. Biol. Chem. 266:4545-4555(1991).
CC -1. SIMILARITY: BELONGS TO THE ABC TRANSPORTER FAMILY.
DR EMBL: M59254; AAA37005.1; -
DR InterPro: IPR003593; AAA_Atpase.
DR InterPro: IPR001140; ABCtransport.
DR InterPro: IPR003439; ABC_transport.
DR Pfam: PF00664; ABC_membrane; 2.
DR Pfam: PF00005; ABC_tran; 2.
DR PRODOM: PD000006; ABC_transport; 2.
DR SMART: SM00382; AAA; 2.
DR PROSITE: PS00211; ABC_TRANSPORTER; 2.
KW ATP-binding; Transport.
SQ SEQUENCE 1169 AA; 128939 MW; 72E25B7CE29DC185 CRC64;

Query Match 82.3%; Score 5333; DB 11; Length 1169;
Best Local Similarity 89.2%; Pred. No. 6,4e-301;
Matches 1043; Conservative 69; Mismatches 57; Indels 0; Gaps 0;

OY 113 MTTTAAVYYSIGAGVLAAYIOVSFWCLAAAGQILKIRKQFHAIMRQEIFGVHDYGE 172
DB 1 MTTTAAVYYSIGAGVLAAYIOVSFWCLAAAGQILKIRKQFHAIMRQEIFGVHDYGE 60
OY 173 LNTRLTDVSKINIEGIDKIGMFQMAITFGGFIIGFRGKRLIVLAISPVGLSAA 232
DB 1 LNTRLTDVSKINIEGIDKIGMFQMAITFGGFIIGFRGKRLIVLAISPVGLSAA 120
OY 233 IMAKILSSFTDKELLYAKAGAAVEVLAIRTVAFGQKLEERYKNLEAKGIGIK 292
DB 1 IMAKILSSFTDKELLYAKAGAAVEVLAIRTVAFGQKLEERYKNLEAKGIGIK 180
OY 121 IMAKILSSFTDKELLYAKAGAAVEVLAIRTVAFGQKLEERYKNLEAKGIGIK 180
DB 1 IMAKILSSFTDKELLYAKAGAAVEVLAIRTVAFGQKLEERYKNLEAKGIGIK 180
OY 293 KAITANTISGAATLLIYASALAFWGTSLVSEKSTGOVLTVEFSVLIGAFSIGQASP 352
DB 1 KAITANTISGAATLLIYASALAFWGTSLVSEKSTGOVLTVEFSVLIGAFSIGQASP 240
OY 181 KAITANTISGAATLLIYASALAFWGTSLVSEKSTGOVLTVEFSVLIGAFSIGQASP 240
DB 1 KAITANTISGAATLLIYASALAFWGTSLVSEKSTGOVLTVEFSVLIGAFSIGQASP 240
OY 353 SIEAFANARGAAYEIFKILNKPISIDYSKSGHPPNIKONLEFKNVHFSYPSRKEVKIL 412
DB 241 SIEAFANARGAAYEIFKILNKPISIDYSKSGHPPNIKONLEFKNVHFSYPSRKEVKIL 300
OY 413 KGLNLKVGSGQTVALVNGSGGKSTVOLMORLYDPIDGVCIDGODITINRHIREIT 472
DB 301 KGLNLKVGSGQTVALVNGSGGKSTVOLMORLYDPIDGVCIDGODITINRHIREIT 360
OY 473 GYVSEPELLEFATTAENIRYGRNVNDELEKAVKEANAYDFIMKLPNKEDILVGRGAQ 532
DB 361 GYVSEPELLEFATTAENIRYGRNVNDELEKAVKEANAYDFIMKLPNKEDILVGRGAQ 420
OY 533 LSGGQKRIATARALVNRPKILLDEATSAIDTESEAVVVALDKARKGTTIVIAHRIS 592
DB 421 LSGGQKRIATARALVNRPKILLDEATSAIDTESEAVVVALDKARKGTTIVIAHRIS 480
OY 593 TVRNADVLAGEDDGYIVKGNHDLMEKGIYFKLYTMOTRGNEIELENTGESKSESDD 652
DB 481 TVRNADVLAGEDDGYIVKGNHDLMEKGIYFKLYTMOTRGNEIELENTGESKSESDD 540
OY 653 LEMSPKDGSSSLIKRRSRRIIAHPOGQDRKLTGKEDLNENVPVSPFWILKLNSTWMPY 712
DB 541 LEMSPKDGSSSLIKRRSRRIIAHPOGQDRKLTGKEDLNENVPVSPFWILKLNSTWMPY 600
OY 713 FVVGIFCALINGLOPAFSIIFSRILGIFTRDEDEPDKRONSMFSLVFLVGLISRTTF 772
DB 713 FVVGIFCALINGLOPAFSIIFSRILGIFTRDEDEPDKRONSMFSLVFLVGLISRTTF 772

DB 601 FVVGIFCALINGLOPAFSIIFSRILGIFTRDEDEPDKRONSMFSLVFLVGLISRTTF 660
OY 773 FLOGFTFEKAGEILTKRLRYVFRSMRLQDVSFDDPKNTGTLTRRLANDAAQVKGATG 832
DB 661 FLOGFTFEKAGEILTKRLRYVFRSMRLQDVSFDDPKNTGTLTRRLANDAAQVKGATG 720
OY 833 SRLAVITONIANLGTGIIISLIYGMOLTLTLLAIVPILAIAGVYENKMLSGQALDKREL 892
DB 721 SRLAVITONIANLGTGIIISLIYGMOLTLTLLAIVPILAIAGVYENKMLSGQALDKREL 780
OY 893 EGAGKIAATEAENRTVSLTRQKFEYMAQSIQVPRNSLKRAHIFGVSEFSTQAMVY 952
DB 781 EGAGKIAATEAENRTVSLTRQKFEYMAQSIQVPRNSLKRAHIFGVSEFSTQAMVY 840
OY 953 ESYACCFRGAIVLANEFMNRQDVLVFAVFGAMVGVSSFPADYAKAKYSAAHVIM 1012
DB 841 ESYACCFRGAIVLANEFMNRQDVLVFAVFGAMVGVSSFPADYAKAKYSAAHVIM 900
OY 1013 IIEKSPILDSYSPHGLKPNTELEGNTEVENYPTRPDIPLVGLSLEVKKQOTLALV 1072
DB 901 IIEKSPILDSYSPHGLKPNTELEGNTEVENYPTRPDIPLVGLSLEVKKQOTLALV 960
OY 1073 SSGCGKSTVVOLEFRYPDLAGSVLIDGKELKHNVMOLRAHIGVSEPTLEFDCSTAE 1132
DB 961 SSGCGKSTVVOLEFRYPDLAGSVLIDGKELKHNVMOLRAHIGVSEPTLEFDCSTAE 1020
OY 1133 IAYGDSRVSHSHEIMQAAKEANIHFEPTLPKXNFRVSGKQTLSSGGQKORIAARAL 1192
DB 1021 IAYGDSRVSHSHEIMQAAKEANIHFEPTLPKXNFRVSGKQTLSSGGQKORIAARAL 1080
OY 1193 VROPHILLDEATSAIDTESEKVOEALDKAREGRTCIYIAHRLSTIONADLIYVPONG 1252
DB 1081 VROPHILLDEATSAIDTESEKVOEALDKAREGRTCIYIAHRLSTIONADLIYVPONG 1140
OY 1253 VKEHGHQOLLAOKGITYFSMVSVQAGAKR 1281
DB 1141 VKEHGHQOLLAOKGITYFSMVSVQAGAKR 1169

RESULT 7
ID 08RA27 PRELIMINARY; PRT; 1275 AA.
AC 08RA27;
DT 01-JUN-2002 (TEMBLrel. 21, Created)
DT 01-JUN-2002 (TEMBLrel. 21, Last sequence update)
DT 01-JUN-2002 (TEMBLrel. 21, Last annotation update)
DE ATP-binding cassette protein B1b.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=SPRAGUE-DAWLEY; TISSUE=LIVER;
RA Yabuuchi H., Ishikawa T.;
RL Submitted (MAR-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL: AY082609; AAJ92458.1; -
KW ATP-binding.
SQ SEQUENCE 1275 AA; 141222 MW; 70D055F6A0E0E19D CRC64;

Query Match 81.6%; Score 5286.5; DB 11; Length 1275;
Best Local Similarity 80.3%; Pred. No. 3.7e-298;
Matches 1029; Conservative 130; Mismatches 114; Indels 9; Gaps 5;

OY 1 MDEPGRRGSAEKNNFKKKSKNEKKKKPYSTFAMRRYNNMLDRYMLVGTAAAT 60
DB 1 MDEPGRRGSAEKNNFKKKSKNEKKKKPYSTFAMRRYNNMLDRYMLVGTAAAT 58
OY 61 HGAALPLMLVAFGMMTSEFANAGISRNKTFPVIIINESITNNTOHFLNHLDEMTYAY 119
DB 59 HGTLLPLMLVAFGMMTSEFANAGISRNKTFPVIIINESITNNTOHFLNHLDEMTYAY 115

```

QY 120 YSGAGVLAAYIYVSWFMCILAAGROILKTRKQFHAIMROEIGMEDVDHVELNRLTD 179
Db 116 YTGAGVLAIVAYIQVSLWCLANGROHKTROKFFHAIMQOIGWPDVADAGELNRLTD 175
QY 180 DVKINGIGDKIGMFHSHATPEPTGFTVGTGKMLTVIIAISPVLGSAIAWAKILS 239
Db 176 DVKINGIGDKIGMFHSHATPEPTGFTVGTGKMLTVIIAISPVLGSAIAWAKILS 235
QY 240 SFTEKELAVAKAGAAVEEVLAAIRTVIAPGCKKLEKRYNNKLEAKIGIKKAITANI 299
Db 236 SFTEKELAVAKAGAAVEEVLAAIRTVIAPGCKKLEKRYNNKLEAKIGIKKAITANI 295
QY 300 SIGAFTLLIYASYALAEWYGTSLVLSSEYSIGVLTFFPSVLIGAFSIGQSPSIEARAN 359
Db 296 SIGAFTLLIYASYALAEWYGTSLVLSSEYSIGVLTFFPSVLIGAFSIGQSPSIEARAN 355
QY 360 ARGAAVEIFKIDNKPSISYSKSGKPNKGNLEFKNVHESYPSRKEVKILKGLNLKY 419
Db 356 ARGAAVEIFKIDNKPSISYSKSGKPNKGNLEFKNVHESYPSRKEVKILKGLNLKY 415
QY 420 QSGQVALVNGSGCKSTVQLMORLYPTDGMVCIDODIRTNVRLRETTGVSOEP 479
Db 416 KSGQVALVNGSGCKSTVQLMORLYPTDGMVCIDODIRTNVRLRETTGVSOEP 475
QY 480 VLEATTIANINIRYGBENVYMDIEKAVAEANAYPEIMKLPKRPDTLVGERGAQSGGOKO 539
Db 476 VLEATTIANINIRYGBENVYMDIEKAVAEANAYPEIMKLPKRPDTLVGERGAQSGGOKO 535
QY 540 RIAIARALVRNPKILLDEATSAIDTSEAVVOVALDKARGRITVIAHRLSTVRNADY 599
Db 536 RIAIARALVRNPKILLDEATSAIDTSEAVVOVALDKARGRITVIAHRLSTVRNADY 595
QY 600 IAGFPDDGVYEGKNDLMEKEKGIYFKLVMTQRTGNEIELEENATGEKSDALEPKD 659
Db 596 IAGFPDDGVYEGKNDLMEKEKGIYFKLVMTQRTGNEIELEENATGEKSDALEPKD 655
QY 660 SSSSLIKRSTRSRSHAPQGGDRKLGKEDENVPVYSWRILKLNSTWPFVVGIFC 719
Db 656 SSSSLIKRSTRSRSHAPQGGDRKLGKEDENVPVYSWRILKLNSTWPFVVGIFC 714
QY 720 AIINGGLQPAFSILFSRIIGITFDEDEPETRKONNSFVLIVLIGISITFELOGFTF 779
Db 715 AVINGCLOPAFPAIVFSKIVGFSRDDHEFKORNCNLSILFLVGMKISVYTFEFGFAF 774
QY 780 GKAGEILKRLRYWVFSMRLQDYWFPDDPKNTGALTTRLANDAAQVAGISRLAVT 839
Db 775 GKAGEILKRLRYWVFSMRLQDYWFPDDPKNTGALTTRLANDAAQVAGISRLAVT 834
QY 840 QNIANLGTGIIISLIYQMOTLILLALVPIIATAGVEMKLSGQALKKKELESGAKTA 899
Db 835 QNIANLGTGIIISLIYQMOTLILLALVPIIATAGVEMKLSGQALKKKELESGAKTA 894
QY 900 TEALENRTVYSLTRECKFEYMTAQSLQVYPYRNSLRKAHIFGVSEFSTQAMMYFSVAGCF 959
Db 895 TEALENRTVYSLTRECKFEYMTAQSLQVYPYRNSLRKAHIFGVSEFSTQAMMYFSVAGCF 954
QY 960 RFAGATVIANEFMFOVLYAFSAIVGAMAYGVSSFADYAKAKVSAHVIMITEKSEPL 1019
Db 955 RFAGATVIANEFMFOVLYAFSAIVGAMAYGVSSFADYAKAKVSAHVIMITEKSEPL 1014
QY 1020 IDSISPGKLPNTLEGNYTFNEVNYPTRPDIPLYQGLSLEVKKGOTLALVSSGGCKS 1079
Db 1015 IDSISPGKLPNTLEGNYTFNEVNYPTRPDIPLYQGLSLEVKKGOTLALVSSGGCKS 1074
QY 1080 TVVOLLERFEDPLAGSVLIDGKEIKHLANQWRAHLHGIYSQEPILFDCSIAENIAYGNS 1139
Db 1075 TVVOLLERFEDPLAGSVLIDGKEIKHLANQWRAHLHGIYSQEPILFDCSIAENIAYGNS 1134
QY 1140 RVVSHHEIQAQAEANIHIFIELEPEKYNTRYGDKGTOLSGGOKORAIATAVROPBIL 1199
Db 1135 RVVSHHEIQAQAEANIHIFIELEPEKYNTRYGDKGTOLSGGOKORAIATAVROPBIL 1194
QY 1200 LDEATSAIDTSEKVVQALDKARGRITVIAHRLSTIONADLIVFQNGKVEHGT 1259

```

```

Db 1195 LDEATSAIDTSEKVVQALDKARGRITVIAHRLSTIONADLIVFQNGKVEHGT 1254
QY 1260 QQLAOKGIFYSMVSVQAKR 1281
Db 1255 QQLAOKGIFYSM-VQAKR 1274

RESULT 8
093437 PRELIMINARY: PRT; 1288 AA.
AC 093437;
DT 01-NOV-1998 (TREMBLrel. 08, Created)
DT 01-DEC-2001 (TREMBLrel. 19, Last sequence update)
DT 01-JUN-2002 (TREMBLrel. 21, Last annotation update)
DE ABC transporter protein.
GN CMDR1.
OS Gallus gallus (Chicken).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;
OC Gallus.
OX NCBL_TaxID-9031;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=INTESTINE;
RX MEDLINE=99209805; PubMed=10195430;
RA Edelman H.M.L., Duchek P., Rosenthal F.E., Foeger N., Glackin C.,
RA Kane S.E., Kuchler K.,
RT "Cmdr1, a chicken P-glycoprotein, confers multidrug resistance and
RT interacts with estradiol."
RL Biol. Chem. 380:231-241(1999).
CC -!- SIMILARITY: BELONGS TO THE ABC TRANSPORTER FAMILY.
DR EMBL: AJ009799; CAA08835.1; -.
DR HSSP: P13569; INBD.
DR InterPro: IPR003593; AAA_ATPase.
DR InterPro: IPR001140; ABCtransportTM.
DR InterPro: IPR003439; ABC_transport.
DR InterPro: IPR000504; RNA_rec_mot.
DR Pfam: PF00664; ABC_membrane; 4.
DR Pfam: PF00005; ABC_tran; 4.
DR ProDom: PD00006; ABC_transport; 2.
DR SMART: SM00382; AAA; 4.
DR PROSITE: PS00211; ABC_TRANSPORTER; 4.
DR PROSITE: PS00030; RRM_RNP_1; UNKNOWN_2.
KW ATP-binding; Transport.
FT CHAIN 2 1288
SQ SEQUENCE 1288 AA; 141917 MW; CB258A5F2826DBEC CRC64;

Query Match 72.0%; Score 4660.5; DB 13; Length 1288;
Best Local Similarity 71.2%; Pred. No. 8; Le-262;
Matches 919; Conservative 160; Mismatches 167; Indels 45; Gaps 12;

QY 2 DPBGKRGSAEKNFWMKSKKNEKEKKP--TVSTFAMFRYSNMLDRLYMLVGTMAAI 60
Db 27 DPDEKKG-----KKKKKKPQVPSPLAFRRSSCTDKLMLFGLSLAIA 71
QY 61 HGAALPLMLVFGNMTDSFANGISRNKTPPYINESTINNTQHTNLHEEMTYAVYY 120
Db 72 HGSISPLAMTIFSDMTDSFTSGMTN-----ITGNSGANSADVFNKLEEMTRYAVYY 126
QY 121 SGIAGVLAAYIYVSWFMCILAAGROILKTRKQFHAIMROEIGMEDVDHVELNRLTD 180
Db 127 SAIAAVALVAAYIYVSWFMCILAAGROILKTRKQFHAIMROEIGMEDVDHVELNRLTD 186
QY 181 VSKINGIGDKIGMFHSHATPEPTGFTVGTGKMLTVIIAISPVLGSAIAWAKILS 240
Db 187 VSKINGIGDKIGMFHSHATPEPTGFTVGTGKMLTVIIAISPVLGSAIAWAKILS 246
QY 241 FTDKELAVAKAGAAVEEVLAAIRTVIAPGCKKLEKRYNNKLEAKIGIKKAITANIS 300
Db 247 FTDKQAAVAKAGAAVEEVLAVRTVIAFGGQEKIKRYHKNLDAKRIGIRKATISNIS 306
QY 301 IGAFTLLIYASYALAEWYGTSLVLSSEYSIGVLTFFPSVLIGAFSIGQSPSIEARANA 360

```

```

Db      :|||||
307 MGAFFLIYASALAEFYGTLLILANLEYSGNVLTVFESVLGASISGOTAFASIAFANA 366
Qy      361 RGAAYEIFKIIDNKPISIDSYSGKHKPDNIKGNLEFRNVHESYSPKREKYLKGLNLYK 420
      |||||
Db      367 RGAAYAIFFNIIDNEPIDSYSDAGHKPDHKGMLNEFONFENFNSPDPDEILKGLNLKN 426
Qy      421 SGOTVALVNSGCGKSTTVOLMORLYDPDGMVCIDGODIRTNVHRLREITGVNSOEFP 480
      |||||
Db      427 CGOTVALVNSGCGKSTTVOLIQREFIDPKEGTTITIDGOLSLNLYKLEIIGVNSOEFP 486
Qy      481 LEATTIAENIRYGRENTYMDIEKAVKANAYDFIMKLPNKFDYLVGERGALSGGQOKR 540
      |||||
Db      487 LEATTIAENIRYGRENTYMDIEKAVKANAYDFIMKLPNKFDYLVGERGALSGGQOKR 546
Qy      541 IAIARALVRNPKILLDEATSALDTESEAVVOVALDKARKGRTTYIAHRLSTYVRNADY 600
      |||||
Db      547 IAIARALVRNPKILLDEATSALDTESEAVVOVALDKARKGRTTYIAHRLSTYVRNADY 606
Qy      601 AGPDGVIVKGNDELKKEKGIYFKLVMTQTRNGNEIELENATGESKSESALEMSPPKDS 660
      |||||
Db      607 AAFESGVITTEQGNHSOLIEKKGIYKLVNMQT-----IETEDPSE--KSEN---AVSYKRS 658
Qy      661 GS-----SL---IKRSTRSIIH---APGQDKRLGTKEPDLNENVPVPFWIILKNSF 708
      |||||
Db      659 GSQSNLDESLKELRGRGSTRSMKKPGEPNDTDEK--GSSPD--BELPPVFLKMLKLNK 715
Qy      709 EMPYVVGIFCAITINGLOPAFSIIFSRIGITREDEPETKRONSNMFSLFVLGIIIS 768
      |||||
Db      716 EMPYVVGIFCAITINGLOPAFSIIFSRIGITREDEPETKRONSNMFSLFVLGIIIS 774
Qy      769 FTFEFLQGFPGKAGEILTKRLMYVFRSMRLQDVSWFDDPKNTTGALTTPFLANDAOVK 828
      |||||
Db      775 FTFEFLQGFPGKAGEILTKRLMYVFRSMRLQDVSWFDDPKNTTGALTTPFLANDAOVK 834
Qy      829 GAIGSLAVITTONINANGIIISLIYGMOLTLILLATVPPIIATAGVYEMKMLSGOALKD 888
      |||||
Db      835 GAGVGRALIAQNTANLGGIIISLIYGMOLTLILLATVPPIIATAGVYEMKMLSGOALKD 894
Qy      889 KKELEGAGKIAETEENFTVSLTREOKFEYVYASLOVPYRNSLRAHIFGVSFSTQ 948
      |||||
Db      895 KIELEAGAKIATEINENTVSLTREOKFEYVYASLOVPYRNSLRAHIFGVSFSTQ 954
Qy      949 AMMYESTYAGCFRGALVLANEFMNODVLLVSAIVFGMAVGOVSPADPYAKAKYSA 1008
      |||||
Db      955 AMMYESTYAGCFRGALVLANEFMNODVLLVSAIVFGMAVGOVSPADPYAKAKYSA 1014
Qy      1009 HVTIIEKSPIDISYSPHGLKPTLEGNVTENEVFNVPTRPDIPYLOGSLEVKKQTL 1068
      |||||
Db      1015 HLEVLEFRVPPIDSYREDGKPEKFGNTRIKDVKNFNPBRPEVKILQGLNLAVEKGETL 1074
Qy      1069 ALVSSGCGKSTVOLLERYDPLASVLDGKETIKHLNWOHLRAHLIGVSOEPLIFDCS 1128
      |||||
Db      1075 ALVSSGCGKSTVOLLERYDPLASVLDGKETIKHLNWOHLRAHLIGVSOEPLIFDCS 1134
Qy      1129 IAEINIAVGDNSRVVSHHEIQAKAEANIHHETLEPEKYNTRYGDKTQLSGGQOKRIAI 1188
      |||||
Db      1135 IAEINIAVGDNSRVVSHHEIQAKAEANIHHETLEPEKYNTRYGDKTQLSGGQOKRIAI 1194
Qy      1189 ARALVRPHILLDEATSALDTESEAVVOVALDKARKGRTTYIAHRLSTIONADIIYVF 1248
      |||||
Db      1195 ARALVRPHILLDEATSALDTESEAVVOVALDKARKGRTTYIAHRLSTIONADIIYVF 1254
Qy      1249 ONGKVEHGTTHOOLLAQGIYFNSVVOAGA 1279
      |||||
Db      1255 ONGKVEHGTTHOOLLAQGIYFNSVVOAGA 1285

```

RESULT 9
 Q91586 PRELIMINARY; PRT: 1287 AA.
 ID Q91586:
 AC Q91586:
 DT 01-NOV-1996 (Tremblrel. 01, Created)

```

DT      01-NOV-1996 (Tremblrel. 01, last sequence update)
DT      01-JUN-2002 (Tremblrel. 21, last annotation update)
DE      Multidrug resistance protein.
GN      XEMDR.
OS      Xenopus laevis (African clawed frog).
OC      Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC      Amphibia; Batrachia; Anura; Mesobatrachia; Pipidea; Pipidae;
OC      Xenopodidae; Xenopus.
OX      NCBI_TaxID=8355;
RN      [1]
RP      SEQUENCE FROM N.A.
RX      MEDLINE=95322451; PubMed=7599185;
RA      Castillo G., Shen H.J., Horwitz S.B.;
RT      "A homologue of the mammalian multidrug resistance gene (mdr) is
      functionally expressed in the intestine of Xenopus laevis.";
RL      Biochem. Biophys. Acta 1262:113-123(1995).
CC      -1- SIMILARITY: BELONGS TO THE ABC TRANSPORTER FAMILY.
DR      EMBL: U17608; AAA75000.1; -.
DR      InterPro: IPR003593; AAA_ATPase.
DR      InterPro: IPR001140; ABCtransport.
DR      InterPro: IPR003439; ABC_transport.
DR      Pfam: PF00064; ABC_membrane_2.
DR      Pfam: PF00005; ABC_tran; 2.
DR      ProDom: PD00006; ABC_transport; 2.
DR      SMART: SM00382; AAA; 2.
DR      ProSITE: PS00211; ABC_TRANSPORTER; 2.
KW      ATP-binding; Transport.
SQ      SEQUENCE 1287 AA; 141505 MW; 06E95073C5771415 CRC64;

Query Match      68.4%; Score 4430; DB 13; Length 1287;
Best local similarity 67.6%; Pred. No. 2e-248;
Matches 871; Conservative 172; Mismatches 211; Indels 34; Gaps 11;

Qy      2 DEGGKSGAENKFNKMKKSKKKEKPEYTFAMFRYSNKLRLYLVTGMAIIT 61
      |||
Db      21 DPNNSK--EKGF--SKPKKKKTEKPKVGVTFMTFRYSSTDKMLFETISLAH 76
Qy      62 GAALPLMLVFGNMTDSFANAG--ISRKTEPVIIINESTINNTQHFINHLEENTYAY 119
      |||||
Db      77 GAALPLMLVFGNMTDSFANAG--ISRKTEPVIIINESTINNTQHFINHLEENTYAY 127
Qy      120 YSGIGGVAAVYIOVSEWCLAAGROILKIRKOPHAIIMROEIGMPDVHVGELNTLTD 179
      |||||
Db      128 YSGIGGVAAVYIOVSEWCLAAGROILKIRKOPHAIIMROEIGMPDVHVGELNTLTD 187
Qy      180 DYSKINEIGDKIYGFHSIAFFTGFIYGFTRGKTLVILATSPYLSAAIMAKILS 239
      |||||
Db      188 DYSKINEIGDKIYGFHSIAFFTGFIYGFTRGKTLVILATSPYLSAAIMAKILS 247
Qy      240 SEFDKELAVAKGAVAEVLAIRTVIAFGQKKELERYNKNLEAKGIGIKKAITANI 299
      |||||
Db      248 AFTNKEIKAVAKGAVAEVLSIRTVIAFGQKKELERYNKNLEAKGIGIKKAITANI 307
Qy      300 SIGAPELLIYASALAEFYGTLLILANLEYSGNVLTVFESVLGASISGOTAFASIAFANA 359
      |||||
Db      308 SIGAPELLIYASALAEFYGTLLILANLEYSGNVLTVFESVLGASISGOTAFASIAFANA 367
Qy      360 ARGAAYEIFKIIDNKPISIDSYSGKHKPDNIKGNLEFRNVHESYSPKREKYLKGLNLYK 419
      |||||
Db      368 ARGAAYEIFKIIDNKPISIDSYSGKHKPDNIKGNLEFRNVHESYSPKREKYLKGLNLYK 427
Qy      420 QSGOTVALVNSGCGKSTTVOLMORLYDPDGMVCIDGODIRTNVHRLREITGVNSOEFP 479
      |||||
Db      428 QSGOTVALVNSGCGKSTTVOLMORLYDPDGMVCIDGODIRTNVHRLREITGVNSOEFP 487
Qy      480 VLEATTIAENIRYGRENTYMDIEKAVKANAYDFIMKLPNKFDYLVGERGALSGGQOKR 539
      |||||
Db      488 VLEATTIAENIRYGRENTYMDIEKAVKANAYDFIMKLPNKFDYLVGERGALSGGQOKR 547
Qy      540 RIAIARALVRNPKILLDEATSALDTESEAVVOVALDKARKGRTTYIAHRLSTYVRNADY 599
      |||||
Db      548 RIAIARALVRNPKILLDEATSALDTESEAVVOVALDKARKGRTTYIAHRLSTYVRNADY 607

```



```

QY 600 IAGDGVIVEKGNHDELMKENGIFKLVMTOT-----RGNEIELENATGESSESDALEM 655
Db 608 IAGDNGVIVEQSGSHKEIMRGVYNNLTQTVETSKTDEDLTHIEYERK-----I 660
QY 656 SPDDSGSLIKRSTRSHIA--PGQDRKLTKE-DLNNENPVPSFMRILKNTEMPY 712
Db 661 PVHTHSNLTNRKSSRNTKSKVPTEDKEVDEEKKKEEGPPVPSFKVAKLNKPEMPY 720
QY 713 FVVGICAIINGIOPAFISIRIIGIFTRDEPETK-RONSNEVSFLVIGIISFIT 771
Db 721 FVVGICAMINGATOPAFIISRIIGVFA---GVQSOMRESSSYSLIFLALGVSFIT 777
QY 772 FFLGCTFEKAGEIILTKRLRYVFSMLRODVSWEDPKNTGALTFRLANDAOGVAT 831
Db 778 FFLGCTFEKAGEIILTKRLRYVFSMLRODVSWEDPKNTGALTFRLANDAOGVAT 837
QY 832 GSRLAVITONIANLTGIIISLYOMOLTLILAVPIIAAGVYEMKMSGALDKRKE 891
Db 838 GTRILALQNVANLTGIIISLYOMOLTLILAVPIIAAGVYEMKMSGALDKRKE 897
QY 892 LEGAKITAEIENFRIVVSLTRQKREYMAQSLQVYRNSLRKAHIFGVFSITQAM 951
Db 898 LEKAKISTDAVNIIRIVVSLTRQKREYMAQSLQVYRNSLRKAHIFGVFSITQAM 957
QY 952 YFSVAGCER-FGAYLIYANEFNFDVLLVFSATYFGAMANGVYSSRPADYAKAKVSAHV 1010
Db 958 VILCOWFSVIGALVVEGLMKLDEVLVSSATVLAGMAGOTSPADPYTKAMISAHV 1017
QY 1011 IMIEKSLIDSYSPHGLKPNTEGNTVNEVYPTRPDIPLYGLSLVKKQGTAL 1070
Db 1018 FSLIERPOIDSYSDQEKCKNCSGNVVEFGVNFNPTRPDIPLYGLSLVKKQGTAL 1077
QY 1071 VGSSCGKSTVYOLLEFVYPLAGSVLIDGKELKHLNVOMLRHAGIYQOEPIIDCSIA 1130
Db 1078 VGSSCGKSTVYOLLEFVYPLAGSVLIDGKELKHLNVOMLRHAGIYQOEPIIDCSIA 1137
QY 1131 ENIAGNSNVVSHHEIMQAKENIHFLETLPEKYNTRYVDKTOUSGGOKOIAIAR 1190
Db 1138 DNIAYGNNRKYQOEIETFAKCANHSFESTIDKYNTRYVDKTOUSGGOKOIAIAR 1197
QY 1191 ALYRPHILLDEATSDALDTESEKYOQALDKAREGRCIYAHRLSTIONADLIVFON 1250
Db 1198 ALIRKRIILLDEATSDALDTESEKYOQALDKAREGRCIYAHRLSTIONADLIVFON 1257
QY 1251 GKVEHGTQOOLAKQGIYFSMVYQAG 1278
Db 1258 GKVEHGTQOOLAKQGIYFSMVYQAG 1285
RESULT 10
090235
AC 090235 PRELIMINARY; PRT; 1348 AA.
DT 01-DEC-2001 (TREMBlrel. 19, Created)
DT 01-DEC-2001 (TREMBlrel. 19, Last sequence update)
DT 01-JUN-2002 (TREMBlrel. 21, Last annotation update)
DE Balle salt export pump.
OS Raja erinacea (little skate).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Chondrichthyes;
OC Elasmobranchii; Squala; Hymnosqualea; Pristiorhina; Batoidae;
OC Rajiformes; Rajidae; Raja.
OX NCBI_TaxID=7782;
RN (1)
RP SEQUENCE FROM N.A.
RX MEDLINE=21340059; PubMed=11447010;
RA Cai S.Y., Wang L., Ballalori N., Boyer J.L.;
RT "Balle salt export pump is highly conserved during vertebrate evolution
RL and its expression is inhibited by pRSC type II mutations.";
RL Am. J. Physiol. Gastrointest. Liver Physiol. 281:G316-G322(2001).
DR EMBL; AF367243; AAK52958.1;
DR InterPro; IPR001140; ABCtransp1TM.
DR InterPro; IPR003439; ABC_transp1.
PFam; PF00664; ABC_membrane; 2.

```

```

DR Pfam: PF00005; ABC_tran; 2.
DR ProDom; PD000006; ABC_transp1; 2.
DR PROSITE; PS00211; ABC_TRANSPORTER; UNKNOWN.1.
SQ SEQUENCE 1348 AA; 148648 MW; 81906F41654CDD CRC64;

```

```

Query Match 51.4%; Score 3331; DB 13; Length 1348;
Best local similarity 49.8%; Pred. No. 1.3e-184;
Matches 661; Conservative 253; Mismatches 355; Indels 58; Gaps 10;

```

```

QY 1 MDEGGRK-GSAEKNFWKGGKSKKN-EKKEKPPVTSFAEFRYSNNILRLVGMMA 58
Db 27 IDGPGGRKYVYNGVYPSKSKDNENLSKDKRIKIGFQLRASCVEYTLWVGVCVA 86
QY 59 ITHGALPLMMVPEGMNDSPANAGISNKKTPPIYINESTNNTOHPIN----- 107
Db 87 LHHGAQRAVLVFLGFLDTFTAYDIELQELKDT--RKHCNNNTIHWNGSEYLNQNER 144
QY 108 -----HLEEMTYYAVYSGIGAGVLYAAVYQVSPFCLAAQRQILKIRQFHALTROE 161
Db 145 MSCGILNTEKEMTMAVYVIGGCVLLIGYFOICFWNTAAARQOTIRAYRQIMRE 204
QY 162 IGMFVHVGELNRLDDVSKIMEGIDKIMGFHSIAFTFTGFTGRKLTLLV 221
Db 205 MGWFDGNSVGLNTRMSDDINRINDAIDQVGIQRFSTFVSGFLMGFVNGKLTLLV 264
QY 222 AISPVLGSAIIMAKIISFTDKELLAYAKAGVAEEVLAIRTYIAFGQKKELERYK 281
Db 265 AVSPLIGIGALMALVAVRLTGLIKAYAKAGVADEVLSSIRTVAFSGEKREVDYDR 324
QY 282 NLEAKGIGIKKATIANISGAFLIYASALAEVYGTSLV--SSVYIGOVLYVFSY 340
Db 325 NLVPAQWQIRKQIMGFETGYWMIIFCCYALAEFYGSKLVLEQNEYRPGLLQVFLCV 384
QY 341 LIGAFSIGQASPSIEFANARCAAYELFRTIDNKPISIDYSKSGKPNIDKNEFKVH 400
Db 385 LVAMNMGQSPOLAEFAFGRGAAVYIFETIDREPIDCSGGLTKLVKGDIEFHVY 444
QY 401 FSVPSREKRIILGLNLKVOGQVVALVNGSGCKSTVQMLQRLYDTPDGAVCIDGDI 460
Db 445 FNPSPRPVKTDLRLSMVIAKGETTAFVPSGSGSTAVELLQRYDYDKQGMVILGDHI 504
QY 461 RTINVHLEITIGVYQOEVLFAITIAENIRYGRNVTMDLEKAVKANAYDEIMKLP 520
Db 505 RSLNIGWMLSLIGIYQOEVLSTIEENIRYRGLKNDIYVAAKANAYDEIMLPE 564
QY 521 KEDTLVGERGAOLSGQORAIARALVNRPKIILLDEATSDALDTESEAVYQVLDKAR 580
Db 565 KENTLVGEGGMSGCGQORAIARALVNRPKIILLDMATSDALDNESEAVYQVLDKAR 624
QY 581 GRTTIVIAHRLSTVNNVINGFDGIVYVEKGNDELMKEKGIYFKLVYQTRGNETELE 640
Db 625 GRTTISIAHRLSTVNNVYITGFEGRAVERGERHAELLERGIYFVLVLOTGGEOLHE 664
QY 641 NATGSKESDALEMPKDSGSLIKRSTRSHIAHAPQDRK----- 683
Db 685 KA---RQVNGAIEDGASEK-RQLIRGSSMASVRSIRHRSSQVSEVLSLSCGDVA 739
QY 684 -----LGTK-----EDLNNENPVPSFMRILKNSTEMPYVVGIFCAIINGIOPAF 730
Db 740 SAVATPISLIDEDKDVDEEESTEPAPVSRILKYVSEMPYMLFSGLSAAVNGCVNPIY 799
QY 731 SIISRIIGIFTRDEDETRKONSMSVFLVIGIISFTFLOGTFGKAGEIILTKRL 790
Db 800 ALFQSILGTFSL-QNEEIKINQINMICFEVVGVLVSLTFOFLOSYFAKAGELTREL 858
QY 791 RYVFRSMLRODVSWEDPKNTGALTFRLANDAOGVAKISRLAVITONIANLTGII 850
Db 859 RKLQFQMLMQLQETIGMDDKRNNSFTLTLRLADASVQAGATQIGMIVNSITNIGVSLI 918
QY 851 ISLYGQMLTLILAVPIIAAGVYEMKMSGALDKRKEGAGKATATAIEFRIVV 910
Db 919 IAFYFSWKLTLLVILCPFLALGALQANMLTGFMQKREALEAGQISSSEALSINRTIA 978

```


Qy	911	SLRREKFEFYMTAQSLQVYPRNSLRKAIIFGVSVSSTIQAMMTSYACGCEFGYLVIANE	970
Dd	979	GLAKREKMFQVLEPAQLEAVYKAIKRAKNAVYGLCEGFQASVFIEMANASYRFGYLVYAEN	1038
Qy	971	MNFQDVLVFSALVFPAMAVGVQSVSPADYKAKAKSAHVMILIEKSPILDIYSPPHGLK	1030
Dd	1039	LHESIVFRTSALVTSGLDGRSSSTTPDYAKAKIAAARFQDLDRPKINIDOSEGEKW	1098
Qy	1031	NTEGATFENEVEFNYPTRPDIIPVLOGLSLEVKKGQTLALVSGSGCKSTVYOLLREFYD	1090
Dd	1099	SNEFGDLIKFECECFEYTPTRDFQVNLNGLSVCVMPQOTLAFGSSSCGCKSISVOLLERFYD	1158
Qy	1091	PLAGSVLIDGKEIKHLNWOCLRAHGLIGVIOEPLDFCSIAENIAIAGDMSRYVSHETIMOQ	1150
Dd	1159	PNQGRVITIDGSTTNVMTAFILRSKIGVIOEPLDSDSIEENIKYGDMSRQVSMDEVIAA	1218
Qy	1151	AKENIHHEFTLEPEKYNTRFVGDKGTQLSGGQKORITAIARALVROPHILLIDENATSALDT	1210
Dd	1219	AKKAQHLHFVMSLPEKYGTRVQAGOSQLSGCKORITAIRALVROPKITILLIDENATSALDT	1278
Qy	1211	ESEKVOEADLDKAREGRTCIYIAHRSTIONADLIVFPNGKVKHEGTHOOLLAKQIGYF	1270
Dd	1279	ESEKIVODALDEARKGRTCIYIAHRSTIONSDILAVMSRGITIIIEGSGHNLMAAKGAYY	1338
Qy	1271	SMVSVOA 1277	
Dd	1339	KLVTTGGA 1345	

[illegible]

Dd	121	DNKRSIDSYSGSKHDPDNIKNGLEKERNVHSETPRKEVYKLIKRLNLKYGSGFTVALVGN	180
Qy	432	GGGKSTTVOLMORLYPTPTGAMVCIGODIRIINVRLREITGVVSOEPLVFTTIAENIR	491
Dd	181	GGGKSTTVOLMORLYPTPTGAMVCIGODIRIINVRLREITGVVSOEPLVFTTIAENIR	240
Qy	492	YGRNVMTDETERKAVKANAAYDFIMKLPKRGDTLYGERGALSGGOKORIALARALVNP	551
Dd	241	YGRNVMTDETERKAVKANAAYDFIMKLPKRGDTLYGERGALSGGOKORIALARALVNP	300
Qy	552	KLILDEATSLDPESEAVVQVALDKARKGRTTIVIAHRLSTVYANADVYAGDDQVYEK	611
Dd	301	KLILDEATSLDPESEAVVQVALDKARKGRTTIVIAHRLSTVYANADVYAGDDQVYEK	360
Qy	612	GNEDELMEKGIYKLYVTMOTRGNELIELENAEGSKSESDLAEMSPKDSGSSLIKRSTR	671
Dd	361	GNEDELMEKGIYKLYVTMOTRGNELIELENAEGSKSESDLAEMSPKDSGSSLIKRSTR	420
Qy	672	RSJHAPOGODRKLGTKEBLENPNVPVSFWRLIKLNSTEMPFVVGIFCAILINGLOPAPS	731
Dd	421	RSJHAPOGODRKLGTKEBLENPNVPVSFWRLIKLNSTEMPFVVGIFCAILINGLOPAPS	480
Qy	732	IIFSRILIGFTDEDEPEFKRONSNMSEVLPVLYGIISFTFPELOGSFFFGKAGELITKRLR	791
Dd	481	IIFSRILIGFTDEDEPEFKRONSNMSEVLPVLYGIISFTFPELOGSFFFGKAGELITKRLR	540
Qy	792	YMFPRSLKODVSWPDDPKRNTTGALLTRRLANDAAYKGAISGRALVTQNTQNTANGTII	851
Dd	541	YMFPRSLKODVSWPDDPKRNTTGALLTRRLANDAAYKGAISGRALVTQNTQNTANGTII	600
Qy	852	SLIYGWOLTLILAIPITIALAGVYENKMLSGOALDKKELEGACKITEAIENTEFPVVS	911
Dd	601	SLIYGWOLTLILAIPITIALAGVYENKMLSGOALDKKELEGACKITEAIENTEFPVVS	660
Qy	912	LTRBOKFE 919	
Dd	661	LTRBOKFE 668	

RESULT	12	
AD	088331	PRELIMINARY; PRT: 1321 AA.
ID	088331	
DT	01-NOV-1998 (TREMBLrel. 08, Created)	
DT	01-NOV-1998 (TREMBLrel. 08, Last sequence update)	
DT	01-JUN-2002 (TREMBLrel. 21, Last annotation update)	
DE	p-glycoprotein sister.	
GN	SPGP.	
OS	Rattus norvegicus (Rat).	
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;	
OC	Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus	
OX	NCBI_TaxId=10116;	
RN	[1]	
RP	SEQUENCE FROM N.A.	
RC	TISSUE=LIVER;	
RA	Childs S.J., Yeh R.L., Hui D., Ling V.;	
RT	"Tax1 resistance mediated by the liver-specific Sister gene of P-	
RT	glycoprotein.";	
RL	Submitted (JUN-1997) to the EMBL/Genbank/DBJ databases.	
CC	-1- SIMILARITY: BELONGS TO THE ABC TRANSPORTER FAMILY.	
CC	EMBL: AF010597; AAC24753.1; "-	
DR	InterPro: IPR003593; AAA_ATPase.	
DR	InterPro: IPR001140; ABC_transportr.	
DR	InterPro: IPR003439; ABC_transportr.	
DR	Pfam: PF00664; ABC_membrane; 2.	
DR	Pfam: PF00005; ABC_tran; 2.	
DR	ProDom: PD000006; ABC_transportr; 2.	
DR	SMART: SM00382; AAA; 2.	
DR	PROSITE: PS00211; ABC_TRANSPORTER; 1.	
DR	ATP-binding; transport.	
SO	SEQUENCE_1321 AA, 146286 MW; 27F67EC36008D0F CRC64;	

SEQUENCE 1294 AA: 142490 MW: DB8ED2602FAA48DF CRC64:
Query Match 44.1%; Score 2857; DB 5; Length 1294;
Best Local Similarity 45.6%; Pred. No. 4.1e-157;
Matches 579; Conservative 239; Mismatches 422; Indels 30; Gaps 9;
22 SKNKKKKKPTVSTFAMFRYSNMIDRLYMGTAIIHGALPLMLVFGMWTSPFN 81
17 SEKKKAPPPKSTIFOLRYRTSTVDRLMAVGIIVSCATGVGJPLMSIIMGVNSQFV 76
82 AGISRRKTPPVIIINESTNNQ-----HFIIHLEEMTYVYVYSGAGVLAAYTQVS 136
77 LG-----TIFLDPNSTASEKAAAEPSHEVIONCKIYY-----LGGIFPANGFLQAS 125
137 FMCILAAGROLKIRKOFPAHIMROEIGMEDVHDELMTRTLTDVSKINIGTKIMFF 196
126 CFMYICKLSNRRRPFHVSVMHQAETAYDKNTSGTLNKLFDNLERVEGTGKVGIAF 185
197 HSIATFPTFTVGTFRGMKLTIVLILASPVIGLSAIVAKILSPFDKELLAVKAGAVA 256
186 QMAAQFICGFVAVATYDMLTLTLMSSLSPPMICGLFLAKLLATATKAEAKQYAVAGGIA 245
257 EEVLAIRITVIAFGCGKLEERYKNLDEKAGIGIKKAITANISGAAPLIIYASYALAF 316
246 BEVLSTIRTVIAFNGOYECKREYEDALEHGKTKIGIKKSFLIGAGLASFEVVIYASYCLAF 305
317 WYGTSLVLSSEYSIGOVLVYFVSVLIGAFSISGOASPSIEAFNARGAAYEIRKIDNKRPS 376
306 WVGTFNVVYSGRLSEGTIVLVEFVSVMGSMALGOAGQOFATIGALGAAASLVEVIDRIPE 365
377 IDSYSKSHKPDNIKGNLEFNHFSYPSRKEVLIKGLNLKAVSQYVYALVNSGCCSKS 436
366 IDATSTEGOTPEKISIGRSIVSKVKEFTYPTRADYVILKGVSLDPOQYALVAGSSCCGGS 425
437 TTVOLMQLRYDPTGMCVDCODIRTNVRLREITGVSECEPLFATTAENIRYGEN 496
426 TTIQLQRYPNDAQOILIDIPLEDENIKYLRQLVGVYVQGEPLNTSIEONIRYGRSD 485
497 VMDELTEKRVKPNANVYFIMKLPNKFDTLVGERGAOLSGGOKORIALARALVNPKITLL 556
486 VSEDIARLAKENANADFKTFPEBLNTLVDRGVQMSGCKORIALARALVNPKITLL 545
557 DEATSAIDTESEAVNOVALDKARKRGTIVIAHRLSTVRNADVAGDDCVIYENHDE 616
546 DEATSAIDTESIESIOSALENARSRGRTIVIAHRLSTVRNADKIYKAKQWMEVGHET 605
617 LMEKGIYFELVYMOFRNGEIELENATGESKESDLEMSPKDSGLIKRRSTR----- 671
606 LIEOKGLIHVLAQVFA---DVDDKPKKKEAERMSRQTSQKRGSVAFKTOESQVDEKP 662
672 RSHAPQODRKLG--TKEDLNENVPVPSFMKILKLNSTEMPYFVVGIFCAIINGLOPA 729
663 GAPAPAPAEKEIKRLKLEEGAVKANLFKILKRAPBEMITIFFALIIALIGAVMFA 722
730 FSIIFSRITIGTREDDEPRTKRONSNMFSLVLVGIISITFFLQGTFTSKAGETILTKR 789
723 FSLFESQIINVS- NPDRODMKKGHPALMFVLAAYGTSMLQCSLFGAAERLIMR 781
790 LRYVFRSMLRQDVSWFDPKNTTGALTITRLANDAAQVAGASRLAVTQVIANLGTGI 849
782 IRSKYRRLVLRQDATYFDMPKHSFGRTITRLATDAPNISALIDYRLGSLFNMAIASVGGGL 841
830 IISLIYGOQLULLLATIAPVPIIAGVYEMKLSGQALKDKKELEGAKIATPAIENRTV 909
842 GIATFYGGOMALVMAIIFPPAAYGQALMKYHGSATSQAKEMENKATMAIAINIRIV 901
910 VSLTRQKFEYMAQSLQVRYNSLRKAHIFGVSFSTQAMTFVSAGGEFGCAVLYV--A 967
902 QALITLQTLKYNIFCSHLDAPHGNGISKAIRIGLTYGPANSIOFTYTAAPFRGLPLIFPK 961
968 NEFNPQOVLVLSAIVGAMAVGOVSFAPDYKAKAVSAAHYIMIIERSPLIDSYSPHG 1027
962 NYLMEPENLVRLVLAISPSFGTIGFASYPPEYIKATFAAGLFLNMLEBPRIDGMTSSG 1021

1028 LKPNTEGNVTEVENVENYTPRPDIPIVQGLSLEVKKGOPLAVGSSCGCKSTVOLLER 1087
1022 TYPO-LSGEVKLKVFYFFRPPAPVILQGLNVHKKPQOTIALVGPSCCKSTVLSLLER 1080
1088 FYDPLAGSVLIDCKETKHNVMOLRAHGLIVSQEPILFDCSIAENIAYGDSNRVSHET 1147
1081 LYPLEGAIVVDNDLQONPKHRLKHTALVSOEPIILFDTISRENIYVIGLOPGETHEQI 1140
1148 MQAAKEANIHFETLPEKYNTRVGDGTQLSGOKORIALARALVROPHTLLDEATSA 1207
1141 ETACSKANIHKFIDELPDGEYFRVGEKGTQLSGOKORIALARALIRPKITLLDEATSA 1200
1208 LDTESEKVOALDKAREGTCIVIAHRLSTIONDLVIFPQNGKVEKHTHQOOLAKG 1267
1201 LDTESEKVOVALDAAKDKTCIVIAHRLSTIYNACIMVWNGOVBEQTHNELIARG 1260
1268 IYFSKAVVOA 1277
1261 AVFALTOKOS 1270
RESULT 14
Q8T9W5 PRELIMINARY; PRT: 1407 AA.
Q8T9W5;
01-JUN-2002 (Tremblrel, 21, Created)
01-JUN-2002 (Tremblrel, 21, Last sequence update)
01-JUN-2002 (Tremblrel, 21, Last annotation update)
DE ABC transporter Abcb2.
GN ABCB2.
OS Dictyostellium discoideum (Slime mold).
OC Eukaryota; Mycetozoa; Dictyostellida; Dictyostellium.
OX NCBI_TaxID=44689;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=AX4;
RA Ajlard C., Loomis W.F.;
RT "Evolution of the ABC transporters of Dictyostellium."
RL Submitted (JAN-2002) to the EMBL/Genbank/DBJ databases.
DR EMBL: AF466305; AAL74249.1;
SQ SEQUENCE 1407 AA: 154763 MW: B1B6DEBER331DAB CRC64:
Query Match 44.0%; Score 2852.5; DB 5; Length 1407;
Best Local Similarity 44.4%; Pred. No. 8.4e-157;
Matches 589; Conservative 250; Mismatches 391; Indels 97; Gaps 14;
20 KSKKNEKKKK--PVSTFAMFRYSNMIDRLYMGTAIIHGALPLMLVFGMNTD 77
107 KKIDEGEKKEGEVGPQVPSFLFRFAKPDILMITIGIGALANGVSPAISIVFGRLMN 166
78 SFANAGISRNKTPPVIIINESTNNQHFIIHLEEMTYVYVYSGAGVLAAYTQVSF 137
167 SESPENLA-DNFDLV--EIVTSNMFYI-----ICGVFVCSYVEAVF 208
138 WCLAGROILKIRKOFPAHIMROEIGMEDVHDELMTRTLTDVSKINIGIDKIGMFFH 197
209 WMLAGEQAVRCRAYIAKAILKOEIIGWYDVTKSELSTRISDPTLLFOGATIEKIGNFLH 268
198 STATFPTPIYFTRG-----WKTLVILIASPVIGLSAIVAKILSPFDKELL 247
269 HTSTICGIVGFVNGKKNYNYLFFCQGLTVLIFALPLIAGAAMTKMADLTKKGD 328
248 AVAKAGAAVEEVLAIKRVIAFGCGKLEERYKNLDEKAGIGIKKAITANISGAAPL 307
329 AVAKAGVAEEKIGSIRVFSFGPEVKRYTERLRLDALDITKGKIMNGIGIGLVFLV 388
308 IYASVLAFWYGTSLVLSSEYSI-----GOVLTYFVSVLIGAFSISGOASPSITAFAN 359
389 LFGYSLSFWSYGKLVDRKKNPVPDRWOGDVLVYFVSIVKGMALGOASPNVASFAN 448
360 ARGAAYEIRKIDNKRPSIDSYSKSHK--PDNIKGNLEKRVNHFYSRKRKVKLIKGLINK 418

```

Db 449 GRGAFKTYEVDNRNKSIDPFSTEGRSLEEFVQGNIEYRNGISYRSRDVAFINNNULT 508
QY 419 VOSGQYALVNGSGCKSTTVOLMORLYDPIDGVCIDGODITINVRHLREITGVASOE 478
Db 509 IKKRTYVALVDGGGKSSVIGLLERFPDPEGEVYLDGTNINIEINISHLRNILVSOE 568
QY 479 PVLFTATIAENIRYGRNENTMDIEKAVKANAYDFIMKLPNNKFDVLVGERGAOLSGGOK 538
Db 569 PVLFRANSIAENIRYGRNENATMDQITLACKTANADFTSALPEGYDPOVEGCVOMSGOK 628
QY 539 QRIATIALYRNKPIILLDEATSALDTESEAVVOVALDKARKGRTTIVIAHRLSTVRNAD 598
Db 629 QRIATIAAKIKDKPILLDEATSALDSONELLVOOSIEKIMIGRTTIVIAHRLSTIODAD 688
QY 599 VIAGFDGVYVEKGNHDELMKKGTYEKLVTMOTRGNIELENAVTESEKSESALDMSGP 658
Db 689 QIAVVGKGAIVETGHELYALNGVYQLVNRQKGD-----DQDKKKKKKSKESKD 742
QY 659 DS-----GSSLIKRSTRRSIHAP-----OGDRKLGTEDELNENVPVPSFWRLTK 704
Db 743 ESNNNIPSSISIDKSIQSIGADSLTSTIGLVNDNNKKKKKKKKKKKKKKKKKKKKKKKK 802
QY 705 LNSTEMRYFYVGIICALLINGLOPAPSTISRTIGITFREDEPTRYKONSNMFVLEFVL 764
Db 803 LSRGDMHFLGLVGLTGLMGAIMPVSTISELIGITQEDQDDELTKRSKRM-ALMFILL 861
QY 765 GIISFTIFLQGFTEGKAGEILTKRLRYVFRSMLRDVSFMDPDKNTTGALTTRILANDA 824
Db 862 AVVALANFIOICYCFTEIGEKTLFMRRLRSPESIMRDIOMFIDLTENSTGRLTAMATEA 921
QY 825 AQVGAIGSRVAVITONINILGTGITISLIGIOMQLTLLLAIPITAIAGVEMKMLSGO 884
Db 922 TLVQGMSTQRLGLLQINITYAGIYAFVSGMKTLTIVLACPVIGFGKTEMDFQGF 981
QY 885 ALKKKELBGNKITEALIENTRYVLSLREOKREFWYASQVPRNSLRKHLRPGSF 944
Db 982 SOKREKVAECGQVASEALIGIRIVSFTCEKILKFRQCDLQPIOMSFKSNVSGLSF 1041
QY 945 SITQAMMYFSGCFRRGAVLVAN-----EFMN----- 972
Db 1042 GFSQCTLFYTLTYLWYGVLDGSEMPAKESTLETYCYNGEYANIGYDEATCIKSFYT 1101
QY 973 ---FQDVLVFSATVAGMAVGOVSSFAADYAKAKVSAHAHYIMIEKSPILDSYSGHLK 1029
Db 1102 TEGSMAARVFFALINSANGVGSQMAFMDLGSKAKLAVALFSLDRVSEIDPFENKQGT 1161
QY 1030 PNTLEGNVTRENVVYPTRPDIPLVGLSLEVKKGOTLALVSSCGCKSTVYOLLERY 1089
Db 1162 LPERKGDIERDKFSTPSRPNKAVFGQNLVIPHGRKVALVNGSGKSSVISTLERPY 1221
QY 1090 DPLAGSVLIDGKIKHLNVQMLRAHLGIVSOEPLIFDCSIAENIAGDNRVVSHEIMQ 1149
Db 1222 NPSGSLTIDGVNIKDLNLMNLMGRNGMGVGOEPFLFSGTIFENIYIGKPDATM--DEYVE 1279
QY 1150 AAKEANIHFFLETLPEKYNTRVGDKGTOLSGOKORAIKRAIARALVROPHILLDEATSALD 1209
Db 1280 AAKAANHTIESLPDAHYTQDCKFTQLSGOKORAIARAIIRNPKULLDEATSALD 1339
QY 1210 TESKEVQVQALAKAREGRCTIYIARBLSTIONADLIIVFONGYKKEHGHQOOLAKGTY 1269
Db 1340 TVSEKVVQVALDVNSKGRISIVIAHRLSTVIDADLIIVYVKEGKVELGTHETTLAENGRY 1399
QY 1270 FSNMVSQ 1276
Db 1400 AELVSRQ 1406

```

```

DT 01-JUN-2002 (Tremblrel. 21, last annotation update)
DE Hypothetical 140.5 kDa protein.
GN C34G6.4.
OS Caenorhabditis elegans.
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabdilita; Rhabditoidea;
OC Rhabdilitae; Peloderinae; Caenorhabditis.
OX NCBI_Taxid=6239;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=BRISTOL N2.
RX MEDLINE=9069613; PubMed=9851916;
RA None.
RT "Genome sequence of the nematode C. elegans: a platform for
RT investigating biology. The C. elegans Sequencing Consortium."
RL Science 282:2012-2018(1998).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=BRISTOL N2.
RX Langston Y., Rohlfing T.;
RA "The sequence of C. elegans cosmid C34G6."
RL Submitted (APR-1997) to the EMBL/GenBank/DBJ databases.
RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN=BRISTOL N2.
RX Waterston R.;
RA "Direct Submission."
RL Submitted (OCT-2001) to the EMBL/GenBank/DBJ databases.
CC -1- SIMILARITY: BELONGS TO THE ABC TRANSPORTER FAMILY.
DR EMBL; U97407; AAB52482.2; -.
DR HSSP; P13569; INBD.
DR InterPro; IPR003593; AAA_ATPase.
DR InterPro; IPR001140; ABCtransport.
DR InterPro; IPR003439; ABCtransport.
DR Pfam; PRF00664; ABC_membrane_3.
DR Pfam; PRF00005; ABC_tran; 3.
DR ProDom; PD000006; ABC_transport; 2.
DR SMART; SM00382; AAA; 3.
DR PROSITE; PS00211; ABC_TRANSPORTER; 3.
KW ATP-binding; Hypothetical protein; Transport.
SQ SEQUENCE 1265 AA; 140465 MW; 4948EF5CA402757 CRC64;
Query Match 45.6%; Score 2827; DB 5; Length 1265;
Best Local Similarity 45.4%; Pred. No. 2,2e-155;
Matches 583; Conservative 246; Mismatches 400; Indels 56; Gaps 14;
QY 21 KSKRNEK-----KEKKPYSTFAFRYSNMLRLVYLVGTMAIIHG 62
Db 2 KSKRNEPTWTKRLKRSSSDSIDESTYKLTNYGILTYTGVDLLDLTGVAVAVHG 61
QY 63 AALPLMLVFGNNTDSFANAGISRNRKTPV-IINESITNTQHFIMHLEEMTYAAYYS 121
Db 62 AGFPLLAIVGGMTVFELRA--QNSDVGVDVWNPGLVPSIDEFNSEYVKKCIYLL 118
QY 122 GIGAGVVAAYIVDSFWCLAGROQLIKTRQFPAIMROGEIGFWDVHDGELNTRLDVDV 181
Db 119 VLGVLMEFTSYVGLAFESYAEVLHKLQNYLKAIIHQDQGFQDQGNLTALTDLDL 178
QY 182 SKIMEGIGDKTIGFHSITFTFTGTYVGTFRGKTLTYIATISPVYGLSAAIMATLSSF 241
Db 179 ERVREGGDGFALLVOMFAALAGVGFYSWMTVMNGFAPLLVLVGAKMSSSMATR 238
QY 242 TDKELLAVAKAGVADEVLAIRTYIAGGOKKLEERYNNLEBEAKGIGIKRAITANISI 301
Db 239 TRVBOETIYAVAGAIAREFTSSIRYVHSLNGHKRLDRFYVALLEVGRQTVIVRYCYMGIGV 298
QY 302 GAAPLLIYASVALAFWYGSIVLSS-EKSIQVLTVEFSVLIGAFSIGQASISAFAPA 360
Db 299 GFSNLCMYSSYALAFWYGSITLIINDPTFRGLITTYFAVALSSSTSLGALRHLSFGTA 358
QY 361 RGAAYELFKTIIDKPSIDYSKSGHKPDNIGNLEFNANHFSTPSKREKYLILKGINLVQ 420
Db 359 RGAASYLRVINSHPKIDYSLDEGLVDNMKGDISFDVHFYPSRKIDTHVLKGISLELK 418

```

Tue Dec 10 07:55:22 2002

us-09-672-725c-2.rspt

Page 14

Qy	421	SGOTVALVNSCCGKSTTVVOLMOGLVDPDQWANCIDGODITIVIRVRLJFETTVSOFEPY	480
Db	419	AGOKIALVSSCCGKSTTVNLQRFDPKRGVLLIDVDLRENVASHLRQIGIVSOFEPY	478
Qy	481	LEPATTIAEMIRIRGRANTYMDIEKAKKEMAAVDITMKLPKPFLLVNEBAOLSGOKOR	540
Db	479	LEFGITTEIKKNGNHNATHDQVBAOCMAANADFKRLPDIGITRVEKVOVLSGOKOR	538
Qy	541	IATARALVBNPKLILDLBATSALDTESEAVOVALDARKGRTTIVIAHRLSTVRNADYI	600
Db	539	IATARALVBNPKLILDLBATSALDTEAREVOGALDQAOGRTTIVIAHRLSTVRNADI	598
Qy	601	AGFDODVIEKENHDELMEKREGIYFKLV-----TWQTRNTELELNATGESKESDOL	653
Db	599	FYVKAGNIYESSHBEIEMSKGIFEDMTQAOVVRQOOQADKLE-----DPISSAHS	652
Qy	654	EMBPKRSGSSLKRRSTRTSRTHAPQOQDRKLTGRKDLANF-NVPPYSWRILKINSTENPY	712
Db	653	HLBRKSTSTSAI---STATSHO-----LAEVECKAPPTSMFKLFKNGDKGW	700
Qy	713	FVVGIGIATINGLOPAPSLIFSNIGITFRDEDEPTRKONSMFSLVFLVIGIISFITE	772
Db	701	FIGIGIGAFIEFSYPAVALVAYEITVNYSLPAD--QKQAVNYFWCMFIMAGITFEYGF	758
Qy	773	FLOGFEFGAGIEILTKRLRYVFPKSLRODVSMEDDPKNTGATTTBLANDAOVGAIG	832
Db	759	FTSANCLGRGCSHLMKTRFEAFANRLLRDIAFYDDLHFGTKLCFTRFADAPNR-VYF	817
Qy	833	SRLAVITONAMLGIIILSLVGNOLZILLAIYPIAINGVEMKLSGOLKXKREL	892
Db	818	TRLPVVLASTVITGALIGTYITGMQALLVYVWPLVLMGCTREKQMRGKQIRTOULL	877
Qy	893	EGAGKATATEIENFTVVSILTRBOKREBYMTAOSLOVYRNSLRKAHLFGVSESTOAMMY	952
Db	878	EERAGKVASQAVHEITVHSLMRBOEPHFYCEYLREPENTMLKAHRYGAVAFSGSLIE	937
Qy	953	FSTVAGCFRGALVANEPMNFODVLVFSATVFGMAAGVSSFAPDYAKAKVSAAHVYM	1012
Db	938	FMYAAAFYLGISTFVNOQAMOPIDVRYEFALISPCGOMIGNTTSLIPVYKARLAASILEY	997
Qy	1013	TIKRSPLIDSYSPHLL-KPNTLEBNVYFNVEYFNPTPPDIPVLOGSLSEVKRGOTLAVY	1071
Db	998	LIEHPPIIDLSDSQIVKPII--GNISIRNVFPNFTPKDKVKTLYGGTTLDIRKAGKVALY	1055
Qy	1072	GSSGCKKSTVVOULERFIDPLAGSVLIDGKEIKHLANWQMLPAHLGIVSÖPILFDCSIAE	1133
Db	1056	GHSGCCSKSTYMLERFYNODRKIMIDGNIRNINISLSREQVCIVSQEPITLDCITGE	1115
Qy	1132	NIAVGNNSVSHHEIMQAAKEANIHIFETLPEKYNRVWDKGTQSGGOKORIIAARA	1191
Db	1116	NICYGTN-RNWTYQCEIVEAAMAHNIHFILGJPDGDTHVHEKSTQSGGOKORIIAARA	1174
Qy	1192	LVNQPPIILDLBATSALDTESEKVOBALDKRREGTCTIVIAHRLSTIONADIVFQNG	1251
Db	1175	LVNPSVLLDLBATSALDTESEKIVGEALDAKQRTCLVIAHRLSTIQNSDVIAYALVSG	1234
Qy	1252	KYKEHGTBOULLAQKGIYSWVSQO 1276	
Db	1235	KIVKCTHDELIRKSEIYOKFCEFTQ 1259	

Search completed: December 9, 2002, 16:58:44
Job time : 1323 secs

• • • •